

Memorial for Hiroshima and Nagasaki

The fiftieth anniversary of the bombing of two Japanese cities should remind the rest of us that, while international control of nuclear energy is no longer feasible, good second-bests are well within our grasp.

THOSE who engineered Trinity, the first real test of a nuclear bomb just over 50 years ago, did indeed change the world as they foresaw (see *Nature* 376, 382; 1995). In the next few days, the tangible outcomes of the Trinity test will be marked by the fiftieth anniversaries of the events at Hiroshima and Nagasaki that announced the arrival of the nuclear age to an earlier generation. The two bombs achieved their immediate purpose: the Pacific War was ended by the Government of Japan, which followed the only sensible course and surrendered. The 250,000 people killed by the two explosions were but a fraction, perhaps 1 per cent, of the civilian lives lost during five years of terrible conflict. What mattered is that they were public proofs of the devastating potential of the new weaponry. For the past 50 years, we have all known that some future conflict might well bring an end to most of our lives.

How have we responded to that knowledge? Well and not so well, as usual. So it is worthwhile to remember, on the plus side, that one of the first responses of the United States was to put forward in the then-new United Nations a plan for the international control of nuclear energy. The Baruch plan (named after the State Department official Bernard Baruch) inevitably had all kinds of defects that might have been removed by negotiation, but it was not given a chance. The Soviet Union dismissed it out of hand as an attempt to create and perpetuate a permanent monopoly for the United States. Two years later, the Soviet Union exploded its own first fission bomb, and then a further decade passed before President Dwight D. Eisenhower raised (in a much diluted form) proposals for international collaboration in nuclear energy whose chief monument now is the International Atomic Energy Agency (IAEA) at Vienna. By then, too many horses had bolted from the stable; China's first bomb, in particular, would have been unstoppable. That was a brave opportunity lost.

Cold War

Instead, we had the Cold War. McGeorge Bundy, who was National Security Adviser at the White House between 1961 and 1966, is right to point to the self-restraint that the nuclear-power states appear to have exercised at times of crisis in the past 50 years. Nuclear weapons were not used during the war over the Korean Peninsula, the Cuban missile crisis or in Vietnam. More importantly, during 50 years, there was no armed conflict of any kind in Western

Europe. There are two ways of describing that state of affairs: 'the balance of terror' and 'nuclear weapons have kept the peace'. The stasis flowing from nuclear deterrence in Europe was nevertheless a mixed blessing. The threat of a nuclear exchange may have kept the peace, but it prevented worldwide opinion from influencing events such as the Soviet crackdown in Hungary in 1956 and in Czechoslovakia in the following decade. So long as there are nuclear powers, those with whom they disagree will be inhibited from protesting.

So why not work towards a world in which there are no longer nuclear powers? The goal is admirable, but there is a sense in which it is logically unattainable. Like the apple in the Garden of Eden, the taste of the forbidden fruit cannot be forgotten. The best that can be done is to work towards arrangements in which most governments appreciate that the nuclear weapons game is not worth the candle, and in which even those who have tasted the forbidden fruit have an incentive to settle disputes by other means than the transfixing threat of nuclear attack.

Achievement

Fifty years after Hiroshima, it is proper to acknowledge how much has been done. Bilateral agreements between Russia and the United States have taken a sizeable fraction of deployed nuclear warheads out of service (but Bundy is right to say that the task of keeping their content of fissile material out of other people's warheads needs more attention than it is being given). The Nuclear Non-Proliferation Treaty (NPT) is also a step in the right direction, especially now that it has been extended indefinitely. The promise of a Comprehensive Test-Ban Treaty (CTBT) next year will help both to cement the NPT and to inhibit the development of new kinds of weapons. But there are also snags.

The problem of unwanted fissile material from Russia is part of a larger problem, that of safeguarding the fissile material (mostly plutonium) arising from the operation of civil nuclear plants throughout the world. The IAEA has the right to inspect most countries' nuclear facilities and to assure itself that fissile material is not being illicitly diverted. But these arrangements are cumbersome and their effectiveness will be further eroded with the passage of time, as the quantities of safeguarded material grow and become more diverse. Yet the statutes of the IAEA permit it to acquire stocks of fissile material, although it does not



have the financial resources (or the political agreement from its members) to do that. Yet with a little ingenuity, it should be possible to devise a role for IAEA as an international broker of fissile material, buying and holding stocks against their use as fuel in civil reactors. Properly designed, such a scheme could even provide established operators with an incentive to send their fissile material for international safekeeping. This opportunity needs to be seized.

Power

A precondition for such a development must be a realistic acknowledgement by governments that civil nuclear power is not the dead duck it now appears to be. Without a future market for fissile material, an internationally managed stockpile would be nothing but a new aid programme writ large — probably very large. But governments have for too long been pusillanimous on this score, chiefly from fear of the voting power of the Green lobby and the calculation that the issue can be safely postponed for a few years. If the civil nuclear industry were recognized for what it is, a means to a safer world, they might (or should) behave differently.

The negotiation of the CBT also needs more thought than it as so far been given. The significance of this step forward is that it would provide the non-nuclear signatories of the NPT with tangible proof that the nuclear powers have moved substantially towards the control of strategic arms. A news report last week from our Paris correspondent (*Nature* **376**, 283–284; 27 July 1995) described the difficulties that have arisen because of the second-thoughts among nuclear weapons states about the definition of the activities that might be allowed under the terms of a CBT. Experiments with sub-critical assemblies of fissile material predate the Trinity test 50 years ago. They entail bringing together a critical mass of fissile material for too short a time for a fully fledged nuclear explosion to take place. The fission energy released during such a test can be arbitrarily small, but revealing to bomb designers.

Sadly, there is no way in which the whole world can be policed to prevent such experiments taking place. The international community should accept that — and the prospect that they will happen. But it should set its face against the construction of facilities in which it might be possible to accommodate nuclear explosions yielding the equivalent of some hundreds of tonnes of TNT. The nuclear powers have taken to saying that they need explosions on that scale to test the safety (and no doubt the continued efficacy) of their weapons. That is a luxury they should be made to do without. Nuclear powers' uncertainty about the value of their stockpiles is one of the comforts a CBT offers the non-nuclear signatories of the NPT.

French ambitions to carry out a series of nuclear tests in the Pacific are said to have been prompted by a wish to gather information before an over-onerous CBT comes into force. It is significant that France is reported to be willing to trade off its testing programme against its demand for an allowed threshold of several hundreds of

tonnes of TNT-equivalent under the proposed CBT. It should be willing to do without either. Indeed, France (like Britain) should now be asking why it has a nuclear weapons programme of any kind. Is Russia now a strategic threat to Western Europe and to France? Would nuclear weapons help in Bosnia and Herzegovina? Does France intend that its nuclear weapons should figure in some equation involving China, and what delivery systems would be appropriate? Neither Britain nor France has a present justification for a nuclear weapons programme. (Britain's argument that its nuclear weapons contribute to NATO's strategic deterrent would be more credible if there were a credible strategic adversary on the European scene.) It would be better for everybody concerned (taxpayers included) if they were to give them up. This is probably one of the few occasions in the past 50 years when such a demonstration of self-restraint would cut some ice with non-nuclear states. If circumstances should drastically change, the forbidden knowledge could always be resurrected.

There remains the strategic question. For the time being, the former Soviet Union is no longer a strategic threat to anybody. Western diffidence on questions such as the membership of Central European states in NATO springs not from the threat of nuclear attack but from the deterrent possibility that Moscow might breathe new life into the plans nurtured by earlier regimes. Sensibly, people have fallen back on diplomacy. China poses other problems. It is natural and inevitable that the most populous country in the world should aspire to the status nuclear power brings. It is also important that China will be, by all accounts, a signatory of the CTB. But China is even less explicit about the reasons why it is a nuclear power than was the Soviet Union in the 1960s. One of the most urgent goals for the years ahead should be to engage China in open and intelligent discussion of this question. Of necessity, only the United States can take on the job — and there is the obvious danger that conversation would lead to the recreation of a Cold War across the Pacific. Far better to aim at agreed self-restraint on nuclear weapons.

Memorial

Years will no doubt pass before all these tasks have been taken up and dealt with. As things are, there is no prospect of an international regime for the management of nuclear power along the lines briefly advertised to the United Nations in 1946. Indeed, so many countries have since then become familiar with the technology of nuclear weapons that the assumption of innocence on which the Baruch plan was predicated is no longer valid. But the next-best regime in which the forbidden knowledge, although widely shared, is hardly used would be better than the present. Irsome intrusions on national sovereignty would no doubt be involved, but so what? Is that not the way in which the world is moving anyway? Would it not be the best memorial for those who died at Hiroshima and Nagasaki that the forbidden knowledge should be voluntarily set aside? □

US immigration moves raise fears over ease of entry for scientists

Washington. University organizations and immigration lawyers are warning that legislation now moving through the US Congress could make it more difficult for foreign scientists and engineers to enter the country. But they disagree about just how difficult.

Their concern has been triggered by a bill introduced into the House of Representatives in June by Lamar Smith (Republican, Texas) that would require some US companies and universities wishing to employ foreign researchers to demonstrate that no American workers are qualified for the job.

Congress is keen to reform the immigration laws during this legislative term, partly in response to public anxiety about illegal immigrants, who compete with American workers and use social services supported by US taxpayers. In addressing such issues, Congress is likely to place new restrictions on legal immigration as well.

Smith's bill, designated H.R.1915, would reduce legal employment-based immigration from its current level of 140,000 a year to 135,000 a year. That alone is not a significant reduction, particularly as fewer than

100,000 such visas are currently granted.

But the bill also changes the categories under which foreigners seeking jobs in the United States can be admitted. In particular, it eliminates the "outstanding researchers" designation whereby high-grade scientists and engineers are exempted from the certification process designed to prove that no US workers are available. Some 1,800 visas were granted under this provision last year.

Legal aliens could escape certification if they have "extraordinary ability in the sciences, arts, education, business, or athletics which has been demonstrated by sustained national or international acclaim". But foreigners with advanced degrees but only "exceptional ability" — and a third tier of "skilled workers and professionals" — would have to follow conventional channels.

Elissa McGovern, a policy analyst with the American Immigration Lawyers Association (AILA), describes the latter as an "elaborate and very long process" that may discourage universities and private companies from hiring foreign researchers. She

also says fewer visas will be available for scientists and engineers, as there will now be in competition in the same category with executives from multinational corporations coming to work in the United States.

But others are less alarmist than AILA, claiming that H.R.1915's category-shifting, while adding some hassle and expense, is unlikely to discourage scientists and engineers coming into the United States.

George Fishman, for example, a staff member on the House immigration subcommittee chaired by Smith, says there is a very high approval rate — more than 90 per cent — for labour certifications. While acknowledging that companies or universities who want to hire non-US researchers may "have more work ahead of them," he believes that in most cases the visas should go through.

John Vaughn of the Association of American Universities (AAU) says that eliminating the 'outstanding researcher' category may pose a problem for universities if the labour certification process becomes so congested that highly-prized faculty members cannot be promised quick entry into the United States. Vaughn says Smith's staff has made it clear he would consider amending the bill to help eliminate this problem.

Smith showed his willingness to bow to university interests last month, when he added an amendment to H.R.1915 that would solve what many believe is a much more serious problem for universities. This was a policy followed by the US Department of Labor (DOL) requiring that foreign researchers, even at postdoctoral level, be paid at industry rates rather than the lower rates normally paid to academic workers.

Acting on a little-publicized ruling by an administrative court earlier this year, the DOL has been cracking down on universities in recent months, insisting that they raise their rates of pay for foreign workers.

But the AAU and other university organizations have asked the DOL to place a moratorium on this practice, arguing that industry and universities represent different kinds of labour market. Smith's amendment to H.R.1915 specifically says that universities would have only to pay rates comparable to those at other universities.

Meanwhile, the Senate is taking its own look at immigration reform. Alan Simpson (Republican, Wyoming), a conservative legislator who has long been an advocate of restricting immigration, now chairs the Senate panel responsible for rewriting the law.

Simpson plans to introduce a bill on legal immigration within the next two weeks that some believe will be more restrictive than the House version. His bill may even set ►

Soliton wave receives crowd of admirers

London. Scientists at Heriot-Watt University in Edinburgh, Scotland's capital city, last month took to the water (right) to honour a Victorian civil engineer for a discovery he made 161 years ago.

John Scott Russell, better known for his successes in ship hull design and the first experimental demonstration of the Doppler effect, is also the first person to have correctly identified a soliton wave, while watching a boat on Edinburgh's Union Canal.

Last month, scientists attending a conference at Heriot-Watt University on nonlinear waves in physics and biology witnessed a reconstruction of that first sighting. The occasion was part of a ceremony to name a new aqueduct after Russell.

A soliton is a nonlinear wave that is able to propagate without spreading out, breaking up or losing strength over distance. It is now the basis of an established method for transporting data through fibre optic cables.

Russell witnessed such a wave while watching a boat being drawn along the



Union Canal by a pair of horses. When the boat stopped, he noticed that water around the vessel surged ahead in the form of a single wave, whose height and speed remained virtually unchanged. Intrigued, Russell pursued the wave on horseback for more than a mile before returning home to reconstruct the event in an experimental tank in his garden.

Until the 1960s, soliton waves were considered a curiosity. But Russell believed he had discovered an important phenomenon, describing the sighting as the happiest day of his life. □

►lower limits on science and engineering professionals coming into the country, or make it more difficult for graduate students to remain in the United States once their education is finished.

Simpson and his staff have been sympathetic to the arguments of young American scientists who complain that a reservoir of foreign graduate students and postdoctoral researchers creates a glut in the market, making it more difficult for US-born scientists to find jobs and keeping salaries down.

Recent reports have supported such argument. David North, a researcher in immigration policy, recently released a study sponsored by the Alfred P. Sloan Foundation called *Soothing the Establishment: The Impact of Foreign-Born Scientists and Engineers on America*. In it he argues that, even though foreign-born researchers are a highly talented group and make a real contribution to US science, their presence in large numbers relieves what he calls the "American Establishment" from spending more resources on recruiting blacks, Hispanics and members of other minority groups.

Simpson and his staff invited North and several other immigration specialists to his office in Washington late last month for an informal briefing, during which the topic of limiting foreign-born scientists and engineers is said to have come up repeatedly. "I was quite surprised at the intensity of interest in restricting high-skilled immigration," says one participant in the meeting.

But even those seeking increased restrictions admit that the United States must be extremely careful in setting any new policy, as it would not want to cut off the supply of talented engineers and scientists.

Once the debate begins, US postdoctoral researchers keen to keep immigration levels down will be opposed by universities and private companies who want the most talented individuals they can find. And in this arena, says one observer, "all the political clout lines up on the side of the universities".

Tony Reichhardt

Embargo system under siege on Wall St over obesity gene

Washington. A \$100-million royalty deal between Amgen, the California-based biotechnology company, and the Rockefeller University in New York, paid off handsomely last week when Amgen's stock surged by 5 per cent in anticipation of results published in the journal *Science* on the effects of obesity gene products on laboratory mice.

But the way news of the results leaked on Wall Street — sweeping away *Science*'s embargo and boosting Amgen's market capitalization by some \$600 million in a day — has raised questions about the status of embargoed information released in advance by journals for use by science journalists.

"There's always a problem when there is unequal access to information," says Teena Lerner of Lehman Brothers, the analyst whose forewarning of three *Science* papers was published in her company's daily client newsletter early on Wednesday, 26 July, triggering the rush on Amgen stock.

Told by reporters that the papers — due for publication on Friday, 28 July — would have to be described in news reports on the share movement, *Science* lifted its embargo on them at 2pm on 26 July. The story led every television news bulletin that night, and its implications have been the talk of this weight-obsessed nation ever since.

In February, Amgen made a down payment of \$20 million, with a promise of up to \$80 million in future royalties, for the exclusive rights to develop products based on Rockefeller's obesity gene work. The agreement followed the announcement by a team of scientists there, led by Jeffrey Friedman and funded by the Howard Hughes Medical Institute, of their successful cloning of an obesity gene (see *Nature* 372, 425; 1994).

The *Science* papers — from groups of scientists at Amgen, Rockefeller and at Hoffman-La Roche at Nutley, New Jersey, respectively — confirmed the effects of injecting a protein product of the gene into mice. Lerner's tip to investors focused not on this result, but on the hype she anticipated would accompany it. "The media's inherent overall interest in obesity and weight loss will likely lead to publicity for these scientific studies on Friday," she told them.

With three research teams involved, and pre-publication information in the hands of 400 science reporters, Lerner defends her action on the grounds that "thousands of people in the United States" knew the papers were coming.

But that knowledge may raise some questions for the Securities and Exchange Commission (SEC), which regulates US stock markets. On Tuesday of last week, any member of the public, for example, could buy 'forward options' to purchase Amgen stock at a later date for 63 cents: on Wednesday, such options were worth \$2.63, and by Thursday, \$4.25.

Insider dealing is by its nature based on stealth, and nothing in the trading record suggests that these options were being bought heavily before Wednesday. An SEC spokesman said that, as a matter of policy, it would not comment on whether an investigation was taking place.

Nan Broadbent, chief of communications at the American Association for the Advancement of Science (AAAS), which publishes *Science*, says that the embargo system would be unaffected by last week's breakdown. "It may be imperfect, but it's the best system we've got," she says.

Colin Macilwain

Developing countries dispute use of figures on climate change impacts

London. An intergovernmental meeting held to finalize a draft document on the social costs of climate change ended in stalemate last week. Representatives from developing countries attending the meeting refused to endorse a suggestion that global warming would cause twice as much economic damage to the industrialized nations as it would to the rest of the world.

Working Group III of the Intergovernmental Panel on Climate Change (IPCC) has been preparing a draft summary for policy-makers of the damage likely to result from a rise in global temperatures after a doubling of carbon dioxide concentrations.

But the drafting ran into controversy when developing nations, led by India, and China, challenged the use of different criteria for measuring damage in countries

of the North and of the South.

The value put on a death in a developed country, for example, was calculated to be 15 times higher than in a less industrialized nation. Such disparities result partly from the conversion of all estimates of loss from national currencies into US dollars. "\$1 in, say, Cambodia is not the same as \$1 in the United States," one delegate remarked.

Also at issue is the value to be placed on the 'abatement costs' of global warming. The IPCC committee had calculated that slowing down global warming could be more expensive than merely paying for the damage caused by a doubling in carbon dioxide concentrations (1.5–2 per cent).

But critics such as Aubrey Mayer of the environmental group Global Commons Institute, based in London, disagree.

Mayer argues that cost-benefit analysis should not be used to assess the damage likely to be caused by global warming. "The difficulties of allowing for risk, or assessing the value of a plant or animal species that becomes extinct, are well known," he says.

Narasimhan Sundaraman, secretary to the IPCC, acknowledges disagreements over putting a value on loss of life. But he adds that industrialized nations' representatives are willing to consider alternative methods of modelling.

At the same time, he points out that developing nations have so far failed to propose a single workable alternative. The IPCC working group will attempt to finalize the policy-makers' summary of its report at its next meeting in Montreal, Canada, in October.

Ehsan Masood

Chemical weapons ban edges to realization

London. Prospects for a worldwide ban on the production, possession and use of chemical weapons edged a step closer last month with the long-awaited publication by the British government of draft legislation to ratify the Chemical Weapons Convention (CWC).

Britain has been one of the key participants in negotiating the convention. Both the United States and Russia, which between them possess most of the world's chemical warfare agents and are said to have used offensive chemical weapons during the past decade, have yet to announce ratification plans (both have signed).

A consultation paper for a bill to ratify the convention, which was signed in Paris in January 1993, was issued by Britain's Department of Trade and Industry on 19 July. The deadline for comments is 29 September and the bill will be debated in the House of Commons when parliament reconvenes in October.

"The Chemical Weapons Convention will help rid the world of these weapons of mass destruction", Anthony Nelson, the trade minister, said in a written statement. "The UK is committed to the convention, and wants to ratify it as soon as possible."

The convention makes it illegal for its 159 signatory states to possess, use or sell chemical weapons. All existing weapons must be destroyed within the first decade of the treaty coming into force. This will take place 180 days after the first 65 countries announce completion of relevant primary legislation in their national parliaments.

The British government's announcement is being widely interpreted as a sign that progress to ratify the CWC should now begin to pick up. Thirty-two countries — including France, Germany, Norway, Spain, Sweden, Switzerland and Australia — have already formally announced completion of legislation to ratify the convention. A further 16 states, including Japan, Canada, South Africa, Algeria and the Netherlands, are understood to be close to ratification.

Competition for the remaining positions is likely to be tough, as only the first 65 nations to ratify will be allowed to influence the body that will be set up to oversee the treaty, the Organization for the Prohibition of Chemical Weapons (OPCW). This agency will employ 470 people at its headquarters in the Dutch capital, The Hague. It will have an annual budget of US\$100 million.

Britain is believed to be keen to join this first tier of countries to ratify the CWC. A British diplomat, Ian Kenyon, already heads the OPCW's international secretariat, based in The Hague, which is a nucleus for the

projected technical secretariat.

Early ratification for Britain, which formally abandoned the offensive use of chemical and biological weapons in the late 1950s, would also secure the long-term future of Britain's historic Chemical and Biological Defence Establishment at Porton Down.

When the United States and Russia announce plans to ratify "the rate at which other countries do so will increase", predicts Julian Perry Robinson, a chemical weapons

IMAGE
UNAMAGABLE
FOR A COPYRIGHT
FOR A COPYRIGHT
REASONS
REASONS

Rex Features

Endangered species? Progress in outlawing chemical weapons (seen here in Iraq) seems set to accelerate.

expert at the Science Policy Research Unit at the University of Sussex.

Countries that have yet even to sign the CWC include Egypt, Iraq, North Korea, Syria, Libya, Taiwan, Somalia and Sudan, as well as those of the former Yugoslavia. All are believed to have used chemical weapons. Israel, another reported user, has signed the CWC, but has yet to ratify.

The British announcement has been welcomed by the country's chemical industry, which will be affected most by the CWC's widespread verification procedures.

The convention is the most rigorous of any comparable international agreements. In addition to banning chemical weapons, it will keep a tight rein on the use of chemicals that can potentially be used in warfare but

also have legitimate industrial uses.

Implementation of the CWC will be the responsibility of a regulatory authority in each country. Organizations dealing in dual-use chemicals will be obliged to declare this information to their respective authorities, and to accept site inspections on demand.

Industry concern had also been heightened because of adverse publicity generated by the association with chemical weapons. "We are disappointed the United Kingdom did not ratify earlier," says David Culpin, director of external affairs at the London-based Chemical Industries Association. It was involved in the framing of the CWC. Industry comments have been taken into account in drafting the government's bill.

But enthusiasm for the draft bill is not universally shared. Some scientists object to the decision not to create an independent agency to oversee the CWC in the United Kingdom. Verification will instead be the responsibility of a unit within the Department of Trade and Industry (DTI).

Doubts about the implications of this decision have been highlighted by the DTI's controversial role in exports to Iraq before the 1991 Gulf War. Critics also point out that the unit will apparently function with little recourse to independent advice, and without a discernible line of accountability.

"Ideally, this unit or organization should have input from independent scientific bodies", says Alastair Hay, senior lecturer in chemical pathology at the University of Leeds. "It is now vital this unit at least be scrutinized by the DTI select committee, so that outside experts and concerned groups can ask relevant and pertinent questions."

The opposition Labour Party, committed to enabling the "speedy passage" of the bill through parliament, stresses its offer is not a blank cheque. "This issue is too important to be left to internal committees", says the party's trade and industry spokesman, Lewis Moonie. "There must be clear accountability to parliament."

Ehsan Masood

Moon landing film 'gave lift' to space station

Washington. The US House of Representatives has given its overwhelming support to a \$2.1-billion budget for work on the international space station in the next financial year, which begins on 1 October, assuring the project's continuation.

In debating an appropriations bill covering the National Aeronautics and Space Administration (NASA) last week, the House rejected by 299 votes to 126 an amendment from David Obey (Democrat, Wisconsin) that would have diverted the funding into social programmes.

It also rejected another amendment from Tim Roemer (Democrat, Indiana), to use the savings for reducing the general budget

deficit, by 287:132. The wide margins reflected strong support for the project from both the administration and the House Republican leadership.

After the vote, Roemer blamed his defeat partly on the success of the movie *Apollo 13*. "I feel like I'm competing against Tom Hanks," he said. But George Brown (California, Democrat), a long-standing supporter of the station who surprised many by defecting to support the Obey amendment, predicted that budget cuts at NASA would eventually make the project unsustainable. "We're cutting the dog's tail off an inch at a time," he said. "You can only go on doing that for so long." C. M.

Canadian research strategy set for lukewarm welcome?

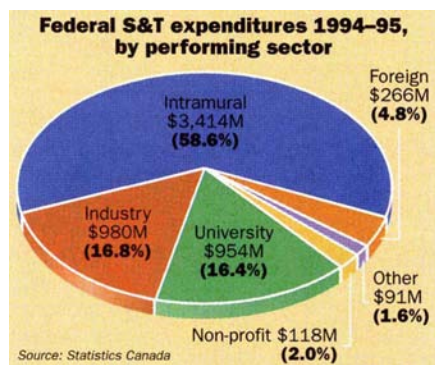
Washington. Plans by Canada's Liberal government to produce a national strategy for science and technology are running late amid allegations that the process, initiated in February 1994, is merely a cover for proposed cuts of up to 20 per cent in the federal science budget.

A draft strategy document prepared for John Manley, the industry minister, was rejected last month by an *ad hoc* cabinet committee, which considered its recommendations too weak to publish. Manley's staff are revising the document, which is now scheduled to be published in the autumn; but many feel that it is still unlikely to contain any radical proposals.

"The focus of the document will be on principles, such as the role of the government in fostering innovation," says one senior government official involved in the process. "It will be impossible for priorities to be set in, say, biotechnology, because so many different departments are involved."

This will disappoint many Canadian scientists, who are keen for the government to adopt a new strategy for science. Such a strategy is needed, they argue, not only to reverse the proposed budget cuts, but also to stem the continuous drain of scientific talent to the United States, and to restructure a system that is skewed against universities and in favour of government laboratories (see pie chart below).

Manley's review process began energetically enough, attracting thousands of scientists and others to a nationwide series of



meetings last autumn. But problems became apparent at the beginning of the year, when the schedule started to slip, and a document that was first promised in early 1995 — and then in June — is now due in October.

According to observers, three main factors have caused the slippage. The first, many claim, is a lack of political leadership from the top of government. Advocates of a radical change want a high-ranking science adviser appointed from outside the government, as well as much stronger cabinet involvement in science and technology. But

Jean Chrétien, the prime minister, has shown little interest in the review since it was announced.

Second, many blame the control exercised over the review process by civil servants at Manley's department, known as Industry Canada. The department, together with associated agencies, controls up to 40 per cent of the government's \$6-billion R&D budget, and its officials are widely considered to be opposed to reforms that could dilute their power.

The third factor is a split within the ruling Liberal party. This is divided over whether its policy for science and technology should be based on a degree of state planning or whether it should be determined by the pressures of the free market.

In April, Chrétien received a report from the National Advisory Board of Science and Technology — made up of leading scientists and industrialists — that called for a cabinet minister to be appointed as a 'champion' of science and technology, for a cabinet committee to deal with such issues and for the appointment of a 'chief science and technology advisor' to the federal government, with his or her own staff.

But the draft document put forward last month rejected the advisory board's proposals, saying that improvements in the running of federally funded science should come from the better management of existing structures. "The government has concluded that the principal issue is not the organizational form of its [science and technology] activities," said the report.

The draft suggested giving existing agencies and departments responsibility for drawing up benchmarks to measure their own success. These efforts would take place against a background of vaguely-worded principles, such as "promoting a stronger science culture" and "capturing the benefits of partnership" with industry.

This was apparently not good enough for the cabinet committee. Nor is it likely to satisfy Canadian scientists. But government officials say that the final document is unlikely to offer much more than the draft version on how science should be run, and the promised national strategy is likely to involve little more than new marching orders for the diffuse array of existing agencies and departments.

These are likely to be told to spend less money, and to do so in ways that boost Canada's economic performance. Action plans from individual departments outlining how these instructions will be met, as well as the final document itself, are now due for publication in early October.

Colin Macilwain

Protestors use World Wide Web in petition against nuclear tests

Tokyo. Using the power of the Internet, three graduate students at Tokyo University have devised a novel and powerful way to protest against French government plans to restart nuclear tests in the South Pacific.

Seishi Shimizu and Yuichi Nishihara, of the university's physics department, and Kiroh Harada, of its faculty of engineering, have set up a home page on the Worldwide Web on which protesters can register their names on a petition that will be delivered to the French Embassy in Tokyo on 4 August.

At the beginning of this week, over 36,000 names had been gathered from 89 countries. Furthermore, the goals and objectives of the petition, originally transmitted in Japanese and English, had been translated into about 20 different languages — including Indonesian, Catalan, Chinese, Dutch, Esperanto, Basque, Finnish, Greek, Polish, Portuguese, Swedish, and Turkish — by some of those who have signed up.

Two of the graduate students started an e-mail chain letter of protest on the Internet on 20 June. But their mailboxes were soon overloaded with thousands of responses arriving each day from around the world. Shimizu had to close down his mail box while Nishihara set up an automatic answering service advising people to switch to the home page. Shimizu says they are "overwhelmed" by the response and are "very sorry" about the problems they caused. But they say that all the e-mail responses have been safely stored, and will be added to the petition.

Shimizu says that the idea was born in response to the fact that the initial reaction of the Japanese public to the planned French nuclear tests has been "very cool". They decided to use the Internet as the "final weapon" to "lower the threshold for action" by individuals both within Japan and around the world.

Of the 36,889 names so far on the petition, 6,961 are from Japan, which is second only to Germany, with 9,983, and names from Japan are pouring in at the rate of about a thousand a day following an appearance of the students on a Japanese television talk show last week.

The Japanese government has already protested strongly to France about the proposed tests. In an unusual move, the Chief Cabinet Secretary, Chews Igarashi, summoned the French ambassador to ask the French government to reverse its decision to resume nuclear tests. ►



Web protest: no French testing!

France thinks small for future satellite plans

Paris. Spurred on by the success of the experimental satellite Topex-Poseidon, France's National Centre for Space Studies (CNES) has announced plans to develop operational satellites that will be similar in design to Topex but four times smaller. But this ambitious programme, called Topex Poseidon Follow-On (TPFO), has yet to receive a promise of government funding.

Launched jointly with the National Aeronautics and Space Administration (NASA) in August 1992, Topex-Poseidon uses altimetric radar combined with a positioning system in order to measure, with great accuracy and almost instantaneously, average sea levels and the altitude of their oscillation.

Orbiting at a height of 1,300 km, the satellite transmits about 500,000 measurements every ten days. Over a period of three years, Topex has allowed scientists to generate an enormous amount of data about both the topography of the ocean surface and the movements and evolution of ocean currents, as well as data about climate patterns.

One of the observations Topex has made possible is that the average sea level has risen by 4 mm a year since 1992. This compares with an average rise of 1.8 mm a year since the start of the century, based on more conventional forms of measurement.

This rapid rise in sea level is probably the result of various phenomena such as the role of the oceans as a heat sink, the expansion of surface water in the oceans and the melting of glaciers and the polar ice-caps.

CNES officials are delighted with the results, and are keen to move on. "Our technological choice has been confirmed beyond our hopes," says Alain Rattier, head of Earth observation at CNES. "Now we want to do with a satellite of 450 kg the same as we did with Topex, with its 2 tonnes."

Speaking during a press conference in Paris last week to announce the new programme, André Lebeau, the president of CNES, talked of a series of such operational mini-satellites which, despite an estimated

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Big brother: French scientists are planning smaller versions of Topex-Poseidon (above).

development cost of a billion French francs (US\$206 million), would be relatively inexpensive while remaining highly reliable.

CNES has also produced a complementary proposal to develop a platform made up of mini-satellites — and provisionally named Proteus — that would be capable of adapta-

NASA tion to various missions.

The main problem for TPFO is who will fund it. Cooperation with NASA, which had been so successful for Topex-Poseidon, is now in doubt. "NASA is very favourable to this idea, but there are problems with the US Congress," points out Rattier (see *Nature* 375, 9; 1995).

Congress will apparently have to choose between supporting Topex or the proposed Geosat Follow-On 2 (GFO-2), which is preferred by the US aerospace industry. "This programme will go ahead with or without the Americans," says Rattier.

Before that, however, the programme must be approved by the government, and then the two parliamentary bodies during discussion of the 1996 budget in the autumn. Even if space is still a protected area of French policy, budgetary restrictions are affecting all sectors of government spending. Financial support for TPFO is therefore far from guaranteed. But CNES officials are optimistic.

Catherine Tastemain

MRC cuts funding of alpha-rated projects

London. Britain's Medical Research Council (MRC) announced this week that the combined effects of financial pressures and high demand from biomedical researchers have made it necessary to reject 70 per cent of grant applications that had received an alpha grading during the review process.

The figures are based on the outcome of decisions over the award last month of £35 million (US\$56 million) in funding for new research programmes. They contrast strongly to the situation in several recent years when the council says it had been able "to fund all or virtually all long-term research proposals in these high quality categories".

As a result of this financial pressure, the MRC has had £5 million a year less to spend over the next four years than for the equivalent exercise last year. This situation has

been partly created by pressure on the council to switch funding to strategic priorities identified through various government programmes, such as the Technology Foresight Exercise and the new ROPAs (Realizing Our Potential Awards) scheme.

"We understand the need for government financial stringency," the chairman of the MRC, Sir David Plaistow, said in a statement. But "no one should underestimate the damaging impact which this will have on the growing opportunities for high quality medical research work relevant to the health and prosperity of the United Kingdom".

MRC officials are careful to point out that the reduced funding available for highly-rated projects is the result of pressures that have built up over several years — similar difficulties are being reported from other research councils — and are, they say, in no way linked to the recent decision to move responsibility for the council's into the Department of Trade and Industry.

Nevertheless, evidence that increased emphasis on strategic priorities may be resulting in reduced funding for investigator-initiated research grants is likely to fuel the continuing debate over the latter move and the government's efforts to harness the science base to the task of wealth creation.

Sir Arnold Wolfendale, for example, the former Astronomer Royal and currently president of the Institute of Physics, wrote in the *Financial Times* at the beginning of this week that such a move was "incompatible" with the desire, even voiced by industrial leaders, to maintain excellence in many scientific disciplines.

David Dickson

► Such a move would conventionally have been made by the Foreign Ministry, and indicates an unusually strong protest by the Japanese government. Furthermore, Makiko Tanaka, head of the Science and Technology Agency, has sent a letter to French president, Jacques Chirac, urging him to reconsider the planned tests.

Until the appearance of the home page on the Web, however, there had not been a strong reaction from the Japanese public. For example, there has been no significant movement to boycott French goods as in some other countries — although one major department store has turned over price labels on French goods to discourage consumers from buying them.

The home page is therefore the first sign

of a significant grassroots movement of protest in Japan. The page declares "stop nuclear tests" and says "we strongly disapprove of the French government's decision to restart nuclear tests". The organizers say they kept the protest simple, and have, for example, not called for a boycott of French goods to encourage the maximum number of people to sign up. But individuals can register their views in an "opinions" corner on the home page.

The organizers will continue to run the home page after they deliver the first petition on 4 August, and will send subsequent updates to the French Embassy. The home page can be accessed at <http://www.icepp.s.u-tokyo.ac.jp/~keshi/>.

David Swinbanks

Shell keeps its options open for disposing of Brent Spar

London. The Brent Spar oil platform could still end up at the bottom of the Atlantic Ocean if an environmentally safer alternative is not found, according to officials from Shell, the Anglo-Dutch oil company.

One month after Greenpeace compelled the oil company to abandon sinking the disused 14,500-tonne North Sea oil platform 150 miles out in the Atlantic, a Shell spokeswoman confirmed that deep-water disposal remains on the agenda.

"Nothing is being ruled out," says Jenny Weller. "We will be going back to the drawing board. We will consider all the available options and then choose the environmentally friendliest."

The British government, which endorsed Shell's decision in October 1994 to pursue deep-water disposal of Brent Spar as the 'best practicable environmental option', says it still supports the idea. Tim Eggar, a minister at the Department of Trade and Industry, said two weeks ago that the company must provide convincing arguments to support any change of plan.

Speaking at a parliamentary briefing, Eggar described the Greenpeace campaign to scupper deep-sea disposal as "dangerous" and "irresponsible", and urged Shell not to change its policy.

The briefing was organized by the government-funded Natural Environment Research Council (NERC), which, along with the University of Aberdeen, has endorsed the deep-sea disposal option. John Krebs, the council's director, maintained that the "available scientific evidence" favoured deep-sea disposal.

The Brent Spar platform, according to figures released by Shell, contains 68,000 tonnes of a concrete ballast chemically similar to rust, 100 tonnes of a bituminous sludge and 30 tonnes of low-level radioactive scale. There are, in addition, small quantities of heavy metals and polychlorinated biphenyls, synthetic molecules similar in structure and action to DDT.

The underwater release of these chemicals, says NERC, poses negligible long-term threats to marine life. The initial impact would affect an area equivalent to two football fields. The organisms affected would be relatively few and would include mainly small organisms such as threadworms, crustaceans and molluscs. "The biological community would be re-established within a few years," predicts Krebs.

Dismantling Brent Spar on land would release environmentally damaging wastes not encountered in the deep-sea option, he says. In addition, the environmental consequences of leaks from the platform or accidental sinking during a towing operation

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

No resting place: deep-sea disposal remains a possibility, backed by the UK government.

would be far greater in shallow waters than in the deep ocean.

But this overall view is still disputed by Greenpeace, whose campaign highlighted the imprecise estimates (also acknowledged by NERC) of the quantities and toxicity of materials contained on the oil platform, as well as a lack of understanding of the deep sea environment.

Paul Johnston, a researcher at the Greenpeace Research Laboratories at the University of Exeter, says he welcomes moves announced by Shell last month for an independent inventory of Brent Spar's contents. But he suggests that a similar, comprehensive assessment be carried out of the impact on marine flora and fauna of the release of chemicals from the platform.

"At present there is very little empirical data on the effects of the proposed disposal on deep sea ecosystems," says Johnston.

Deep-sea disposal of Brent Spar, may lead to similar disposal of 50 deep-water oil installations in UK offshore waters that are next in line for decommissioning, Johnston says. "No one knows what effect this would have on the marine environment."

Greenpeace favours land-based decommissioning of oil rigs, explains Johnston, as facilities and methods for disposing, recycling and de-toxifying oil-industry waste are well-established in many parts of the world.

Brent Spar is now anchored at Erfjord in Norway, awaiting Shell's next move, which is likely to be later this year. Shell has, in the meantime, accepted Greenpeace's offer, along with that of "other bodies and environmental groups", to help develop an alternative disposal plan.

Ehsan Masood

Boost for Japanese centres of excellence in new grant scheme

Rex Features

Tokyo. Japan's Ministry of Education, Science and Culture (Monbusho) last week announced the first winners of substantial research grants under a new scheme intended to cultivate national centres of excellence within Japan's universities.

Grants worth about 300 million yen (over \$3 million) a year for five years were awarded to groups of researchers at Tokyo, Kyoto, Osaka, Nagoya and Hitotsubashi universities, and at Tokyo Institute of Technology. The six were chosen after a strongly contested competition that had attracted 194 applications.

The grants mark a historic departure from the ministry's traditional approach of spreading grant money relatively evenly — and thinly — throughout the university system. They also indicate where some of the best research in Japan is being carried out.

The value of each grant greatly exceeds that of the ministry's 'special distinguished' and 'priority' grants, up to now Monbusho's most coveted awards. All have been awarded to teams working on research in science and technology, except for one on economics at Hitotsubashi University.

A team of about 60 researchers at Kyoto University, for example, headed by Shigetada Nakanishi of the Institute of Immunology, has been awarded 280 million yen in its first year to explore principles and molecular mechanisms of higher-order biological systems. The team, which includes a large number of graduate students, plans combining study of the nervous and immune systems with research on development and intracellular signal transduction mechanisms.

Another team made up of nine faculty members at Tokyo University, and headed by Katsuhiko Sato of the physics department, will receive 320 million yen to study the origin of the Universe. Ryoji Noyori of the chemistry department of Nagoya University will lead a team including 10 faculty members, and over 40 graduate students and postdoctoral fellows, in the synthesis and study of chiral substances.

Kenichi Iga, director of the Precision and Intelligence Laboratory at Tokyo Institute of Technology, plans to develop novel parallel optoelectronic systems with a team of nearly forty faculty members, graduate students and postdoctoral fellows.

Finally, another team of researchers led by Kenji Sobue of Osaka University medical school has been awarded support to pursue the ambitious goal of elucidating the whole picture of signal transduction both within and outside cells. The team's aim is to understand the pathogenesis of and discover new treatments for diseases such as cancer, AIDS, and arthritis.

David Swinbanks

Figures confirm global R&D spending trends

London. Total spending on research and development (R&D) by the world's industrialized nations fell in 1993 for the first time in many years, apparently as a combined result of a drop in military research spending by governments, and lower research budgets in private industry.

Indeed, partly as a result of the economic recession hitting the private sector, and despite calls for a greater market-oriented approach, governments' share of total national research expenditure has been increasing in several countries. In Japan, for example, it rose from 18.0 to 21.4 per cent of gross expenditure on research and development (GERD) between 1990 and 1993.

But if research expenditures are being cut back, there has been a dramatic increase in efforts to secure foreign markets for the results of research as indicated by patent applications — one of a set of research 'output' measurements that are becoming increasingly popular in science policy circles.

In particular, the overall number of 'external' patent applications, defined as patent applications filed abroad, and used as a measure of the rate at which technology is being diffused internationally, doubled between 1987 to 1992 (a period in which 'resident' applications, lodged in the countries in which inventions were made, grew by only 13 per cent).

The United States was particularly active in patenting abroad, with external applications growing from 176,763 to 413,439, an increase of 247 per cent. Growth in Britain was slightly less, but still high, almost doubling over the five years. Both Germany and Japan, with long traditions of foreign patenting, increased such activity by 45 per cent.

All these figures are contained in the latest statistics on science and technology which have been drawn together from a variety of national and international sources by the Paris-based Organization for Economic Co-operation and Development (OECD).

The United States still heads the list of research spenders. This is true both in absolute terms (its domestic expenditure of

\$170.0 billion on R&D and 1993 represents 44 per cent of research spending by all 24 member states), and in terms of spending per head of the population (see figure right).

As far as research spending as a proportion of gross domestic product is concerned, however, it is a different story. Here, Japan still dominates the larger countries, at 2.93 per cent, even though this has fallen from 3.08 per cent in 1990 (see table below).

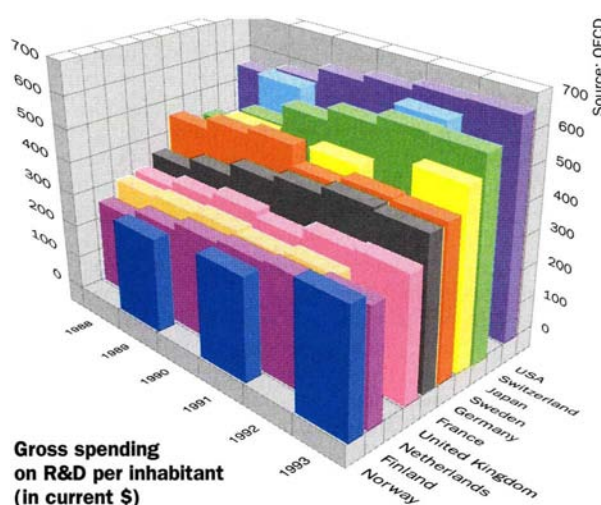
Furthermore, although research spending by the United States on this measure is still greater than most of its European competitors, it is still less than some smaller countries, such as Sweden at 3.12 per cent.

Another equally significant trend to emerge from the OECD statistics is the worldwide decline that took place at the beginning of the 1990s in industrial support for research. As officials in the government of British prime minister John Major have been proudly pointing out, the United Kingdom was the only large industrialized country in which industry actually increased its support of R&D (by an annual rate of 4.6 per cent) in 1993.

In contrast, and for a variety of different reasons (mainly linked to the economic recession), such spending was on decline in all Britain's major competitors, at rates ranging from 0.8 per cent in the United States to 6.6 per cent in Japan. This global trend was a major contributor to an overall drop of 0.7 per cent in R&D spending as a whole by OECD member states.

In addition to compiling statistics on research spending, analysts at the OECD are now turning increasing attention to output measurements widely used to indicate the effectiveness of investment in R&D.

Patent statistics, for example, provide a



more optimistic picture than research budgets alone. Overall, the number of patent applications in the 24 states rose by 47 per cent between 1987 and 1993,

However, particularly large increases in smaller countries such as Denmark (354 per cent) and Greece (280 per cent), reflect the growing interest of foreign companies in the markets of such countries as much as any growth in indigenous patentable research.

A further indication of this trend is given by the fact that whereas the 'dependency ratio' — the ratio of non-resident to resident patent applications — increased only from 0.95 to 1.01 over these five years in the United States, it grew from 2.28 to 4.00 in the member states of the European Union over the same period.

Even so, the United States did relatively badly in terms of 'inventiveness coefficient', the number of resident patent applications per 10,000 population. Top of the list came Japan, a highly patent-conscious society, with a ratio of 27.2.

Second was Switzerland, with a ratio of 4.8, and third — despite its low spending on R&D — was Australia, with 4.5. By contrast, the United States came seventh, below Germany, Finland and Sweden, and just in front of the United Kingdom. □

Country	R&D spend \$ billion 1993	Annual growth in R&D spend 1990 1993	National R&D spend as % of GDP 1990 1993	Government spend as % of R&D 1990 1993	Civilian spend as % of govt R&D 1990 1993	Rate of increase in industry R&D 1990 1993
USA	170.0	3.2% -0.5%	2.82% 2.72%	43.8% 39.0%	37.4% 41.0%	6.6% -0.8%
Japan	74.8	8.3% -2.4%	3.08% 2.93%	18.0% 21.4%	94.6% 93.9%	10.0% -6.6%
Germany*	37.2	1.5% -1.1%	2.76% 2.48%	34.1% 37.1%	86.5% 91.5%	1.2% -2.4%
France	26.0	6.1% -0.8%	2.41% 2.41%	48.3% n/a	60.0% 66.5%	5.1% n/a
UK	21.6	1.9% 2.5%	2.23% 2.19%	34.8% 32.3%	57.5% 54.9%	-0.3% 4.6%
Italy	13.2	6.7% -1.3%	1.30% 1.30%	51.5% 45.9%	93.9% 93.5%	0.6% -4.4%
N America	180.2	3.3% -0.4%	2.70% 2.44%	43.8% 39.6%	39.9% 44.6%	6.5% -1.0%
EU total	123.1	3.7% -0.3%	2.00% 1.96%	40.6% n/a	77.2% 81.1%	2.1% n/a
OECD total	385.5	4.3% -0.7%	2.39% 2.24%	37.7% 36.2%	60.2% 64.1%	5.8% -2.2%

*figures for Germany cover western Germany in 1990, and unified Germany in 1993.

Selected statistics on R&D spending in OECD countries over the period 1990–1993.

Source: Main science and technology indicators, OECD

Science in Australia

SIR — Your appraisal of science in Australia (*Nature* 375, 177–182; 1995) concluding “looks up and ahead” can only be the truth, given the sorry state of support for science, as viewed from the bench, over the past decade or more. The quality of Australian science remains remarkably high given the infrastructure and resources available and is a credit to scientists still working in that country.

You point out that Australian salaries are low by world standards yet science in Australia is more truly “international” than that elsewhere, and even that members of the Waite Institute in Adelaide appear to relate to Europe rather than Sydney. You are of course right. The reason is that many of the best scientists have to leave to pursue their career goals.

The drain to Europe and North America risks becoming a flood, and recent trends, such as the virtually annual reviews of CSIRO, crowding at the universities and the perennial Australian ‘tall-poppy syndrome’ will not slow it down. Australian scientists are replaced by temporary postdoctoral workers and other, largely British, academics seeking to enjoy the life ‘down under’.

Australian science is impressive and has good reason to look up and ahead, but if all expatriate Australian scientists were still there with the resources they now enjoy elsewhere, what a story that would be.

Craig M. Liddell

New Mexico State University,
Department of Entomology, Plant
Pathology and Weed Science,
Box 30003/Dept 3BE,
Las Cruces,
New Mexico 88003-0003, USA

Cerebral malaria

SIR — We have found that most of the clinical isolates of *Plasmodium falciparum* from the recent epidemic of cerebral malaria in Rajasthan, India, were resistant to the most commonly used drug, chloroquine (21 out of 22). But we also found that all tested isolates among these resistant cases were sensitive to mefloquine (18 out of 18) as well as to quinine (10 out of 10). Most of the clinical isolates were also sensitive to pyrimethamine/sulfadoxine (9 out of 10). Alternative antimalarial drugs were therefore still available, if anybody wished to use them.

Molecular studies (using PCR) based on knob protein gene show the presence of multiple strains of *P. falciparum* in these chloroquine-resistant isolates. We have found at least three distinct alleles (thus parasite strains) of this gene — one rare

type, one common type (both reported earlier¹) and one intermediate between these two. Some patients were even harvesting the mixture of both alleles (rare and common type). The finding of an intermediary chloroquine-resistant strain in one individual and a mixture of two strains (also resistant to chloroquine) in other patients may be an indication that the newer but chloroquine-resistant strains were emerging quickly during the epidemic. The emergence of chloroquine resistance seems to have resulted from doctors or state health workers prescribing 600 mg chloroquine to malaria patients irrespective of species. This is certainly a low dose for *P. falciparum* malaria².

The present epidemic is a result of poor management of malaria control programmes, as has been pointed out³. So it seems that not only has the wrong drug policy resulted in the emergence of multiple chloroquine-resistant strains but there has also been a high death toll because of the inability to detect them in time and to treat the patients with more appropriate antimalarial drugs such as mefloquine or quinine or even pyrimethamine. If it is true that some of these antimalarial drugs (for example mefloquine) are not available in India for such treatment, is it not time to review the policy?

Y. D. Sharma

Rajiv Kant

Department of Biotechnology,
All India Institute of Medical Sciences,
New Delhi 110 029, India

C. R. Pillai

M. A. Ansari

Usha Fillai

Malaria Research Centre,
22 Sham Nath Marg,
Delhi 110 054, India

1. Sharma, Y. D. *Nature* 373, 279 (1995).

2. A Note on Malaria Control and Prevention in Rajasthan (Medical and Health Service, Government of Rajasthan, Jaipur, 1994).

3. Sharma, V. P. *Nature* 369, 700 (1994).

Apoptosis again

SIR — The word “apoptosis” was taken straight from Liddell and Scott’s classical Greek-English lexicon complete with examples of its use in medicine by Hippocrates and Dioscorides (the physician, not the poet). It should therefore be pronounced with two ‘p’s and the accent on the antepenult. In a footnote to the landmark paper (*Br. J. Cancer* 26, 239–257; 1972) in which they introduced the term into modern science, J. F. R. Kerr, A. H. Wyllie and A. R. Currie give apoptosis in Greek letters but beginning with the wrong breathing and ending with the wrong sigma, and they suggest silencing the second p and accenting the penult because they wanted a word that would chime with mitosis: hence the present controversy.

David Latchman (*Nature* 375, 100; 1995) very properly prefers the observance of strictly defined criteria for the recognition of apoptosis to arguments about its pronunciation. Indeed, matters of definition and terminology go hand in hand. A consensus statement on apoptosis with 29 signatories (A. Alles *et al.* *FASEB J.* 5, 2127–2128; 1991) condemned the use of the terms “programmed cell death” and “cell suicide”, but even some of the signatories failed to take much notice. The term ‘developmentally programmed cell death’ is acceptable because development is a sort of programme, but programmed cell death as an alternative term for apoptosis is not, unless a programme is considered to be any sequence of cause and effect, in which case, all deaths, including necrosis, would be covered. Most deaths involve mechanisms that are determined by the anatomy, physiology and chemistry of the thing dying; but if patient deaths are to be certified as ‘programmed’ or ‘suicide’, because genes play a major role, there may well be problems with the coroner.

Although many would object to an arbitrary change in the pronunciation of apoptosis, it is certainly necessary to give the term a closely defined meaning for modern scientific use. Naturally Liddell and Scott are silent on this topic, but as Humpty Dumpty (both ‘p’s pronounced) stressed to Liddell’s daughter Alice “A word means just what I choose it to mean — neither more nor less. The question is which is to be master — that’s all.”

W. Beautyman

Beautyman Associates,
1201 Bethlehem Pike,
Flourtown, Pennsylvania 19031, USA

Bats galore

SIR — It is remarkable that echolocation in bats was treated as a new finding 50 years ago (“50 Years Ago”, *Nature* 1 June 1995, page ix and 155, 672; 1945). A quarter of a century earlier, Hamilton Hartridge presented convincing arguments that bats used this method to avoid obstacles: in this case, threads hung across a room (*J. Physiol., Lond.* 54, 54–57; 1920). Hartridge (1886–1976) was one of Britain’s most versatile physiologists, and is best remembered for his early work on haemoglobin and his later, rather idiosyncratic, views on human colour vision.

It is sobering to reflect that as late as 1919 Hartridge could obtain “sometimes between one and two hundred [bats] at a time” in his room at King’s College, Cambridge, by leaving the windows open on summer nights.

C. R. Cavanaugh

Institut für Arbeitsphysiologie,
Ardeystrasse 67,
D-44139 Dortmund, Germany

A road by any other name

SIR — I agree with M. Benarie (*Nature* 376, 12; 1995) that famous British scientists are not as well known in Britain as are native scientists in other countries, but in London alone there are 16 roads named after Newton, 10 after Harvey, 6 after Haldane and 3 after Dalton. Even Bessemer and Flamsteed have been honoured. And Pasteur appears as a road name twice in the capital.

Karl J. Hunter

Flat 4, 11 Shakespeare Road,
Bedford MK40 2DZ, UK

SIR — The Malpas area of Newport, Gwent, Wales, has Baird Close, Chadwick Close, Darwin Drive, Faraday Close, Haldane Place, Kelvin Close, Lister Green, Newton Way, Rutherford Hill, Trevithick Close and Whittle Drive, to give just a selection.

John Twin

The Patent Office,
6 Conde Close,
Ross-on-Wye,
Herefordshire HR9 5UR, UK

SIR — There is a Darwin Drive in Cambridge. Other roads in this city named after famous scientists include Cockcroft Place, Dalton Square, Davy Road, Newton Road, Rayleigh Close, Rutherford Road and Harvey Road.

P.-L. Chau

Department of Pharmacology,
University of Cambridge,
Cambridge CB2 1QJ, UK

SIR — In Ipswich (a town not noted for its academic links and unusual in that it still has no university), Lister, Kelvin, Rayleigh and Sherrington all have streets named after them and are within 400 metres of each other. Newton Street and Newton Road are in other parts of the town. In the humanities, Shakespeare, Arnold, Blake, Browning, Bunyan, Burns, Byron, Carlyle, Chaucer, Coleridge, Constable, Crabbe, Defoe, Dryden, Gainsborough, Garrick, Goldsmith, Kingsley, Kipling, Macaulay, Mitford, Thackeray and Wordsworth are also honoured.

John McCauley

29 Waverley Road,
Reading, Berkshire RG30 2QB, UK

SIR — In Bologna, where I was born, there is not only a via Isacco Newton but also a via William Shakespeare, a via Giorgio Byron and a via Carlo Dickens. Other streets, piazzas and avenues are named after other famous men of letters, as well as musicians, artists and other personalities of all sorts. Among scientists, there are streets named after Einstein, Copernicus, Galileo, Pasteur, Marconi and Fermi.

In contrast, the toponymy of British and American cities seems bland and sterile. Britain and the United States have no dearth of great men and women — they should have their place in the sun.

Cesare Emiliani

Department of Geological Sciences,
University of Miami,
Coral Gables, Florida 33124, USA

SIR — M. Benarie, in trying to persuade the British to abandon the healthy habit of *not* naming streets after eminent people, quotes the Italian example. This is hardly commendable. We have had in recent history at least five waves of street name changes.

The Risorgimento and unity produced a vast shift from saints to Savoia royalty and Garibaldian characters, followed by re-namings to make room for the heroes and places of the First World War. Then came the turn of the Fascist martyrs and the like. The Second World War (and our skill in ending it as allies of our former enemies) and the civil war in northern Italy produced such a turmoil of changes as to make the average Italian envious of the impersonal stability of the English street.

Marco Fraccaro

College Cairoli (formerly
Germanico-Ungarico),
27100 Pavia, Italy

Birth dates

SIR — Recent correspondence¹⁻³ about the effects of birth dates on success in various sports prompted us to investigate birth dates of students entering Porto faculty of medicine in the past few years. People choose to study medicine for various different reasons, but in recent years students entering Portuguese medical schools have had one thing in common: they are among the most successful high school students. This is due to the scarcity of places in medical school compared with the number of interested candidates, and has resulted in the exclusion of any student with marks of less than 16–17 out of 20 in a series of cognitive tests.

Our analysis revealed an asymmetry of birth-date distribution, constant during the three-year period of study. The dates of birth of 263 medical students, compiled into quarters, are shown in the table ($\chi^2=16.437$, $P<0.001$).

As can be seen, there is a significantly higher incidence of students born during the second trimester of the year. This distribution does not reflect any asymmetric distribution of births in Portugal during the years of 1973–75 when the students were born (detailed figures available from

the authors on request).

A relative age advantage cannot be invoked to explain these data, as is sometimes done to explain success in sports¹⁻³. We have no clear-cut answer, but the hypothesis considered self-evident by Michael Holmes⁴ to explain the relation-

PORTO FACULTY OF MEDICINE STUDENTS IN
BIRTH-DATE QUARTERS

Year of entry	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec
1991-92	19	34	19	20
1992-93	17	30	20	21
1993-94	17	30	19	17
Total	53	94	58	58

ship between season of birth and stance taken in two scientific revolutions by prominent scientists seems to be worth attention: seasons are experienced at different stages of early development, either indirectly through maternal behaviour and biology before birth, or directly in the first year of extrauterine life. During neurobiological development, "the brain must wire itself"⁵, being very much dependent on external stimuli during critical periods of early life, so that kind of proposal makes real sense.

Isabel Azevedo
Perpétua Pinto-do-Ó
Nuno Borges

Faculty of Medicine,
University of Porto,
4200 Porto,
Portugal

1. Dudink A. *Nature* 368, 592 (1994).
2. Edwards, S. *Nature* 370, 186 (1994).
3. Baxter-Jones A. & Helms, P. *Nature* 370, 186 (1994).
4. Holmes, M. *Nature* 373, 468 (1995).
5. Stryker, M. P. *Nature* 375, 277–278 (1995).

Down House

SIR — "Down House — the birthplace of Darwin's theory of evolution by natural selection"¹. Childhood home of the theory, maybe, but birthplace, no. The complete theory was surely born in late September 1838 on Marlborough Street in London, with the famous reading of Malthus, and the realization of the importance of competition among members of the same species^{2,3}. The Darwins did not see Down House until July 1842, and did not move into it until September 1842 (ref. 3) — four years after the theory's birth, and three months after Darwin had written out a brief rough draft of the theory.

Alexander H. Harcourt

Department of Anthropology,
University of California,
Davis, California 95616-8522, USA

1. *Nature* 375, 546 (1995).
2. Mayr, E. *One Long Argument. Charles Darwin and the Genesis of Modern Evolutionary Thought* (Harvard University Press, Cambridge, 1991).
3. Desmond, A. & Moore, J. *Darwin* (Michael Joseph, London, 1991).

Fifty years in the shadow of the bomb

McGeorge Bundy

The world is still sharply divided over Hiroshima and Nagasaki. But we have since created a tradition of non-use of nuclear weapons that is a much better guide for the future than were those fateful bombings 50 years ago.

THIS summer we are remembering the bombings of Hiroshima and Nagasaki. These events fully deserve remembrance, but their fiftieth anniversary next week should also lead us to reflect on the extraordinary fact that what happened then has not happened since. The bombings themselves have a double meaning, of deliverance and destruction, that can still sharply divide people. But there is no such division over the subsequent 50-year worldwide record of non-use of nuclear weapons. We may attribute the result to different causes: this or that state has had effective deterrent strength; leaders with the power to fire weapons have shown statesman-like restraint; public opinion, at home or abroad or both, has held them back; or mankind has been lucky. These varied interpretations have claims on our attention, but it is right to begin with the recognition that the division over the attacks on Hiroshima and Nagasaki has no parallel when it comes to the record of non-use in the succeeding half-century; we are all glad about that achievement.

What makes for hot debate about the two bombs dropped in August 1945 is that people give differently weighted answers to the question of what would have happened without them. One side — among Americans the majority side — still emphasizes that the use of the new weapon ended an appalling war that desperately needed ending. It can cite the detailed record of the rapid political interaction set off by the Hiroshima bomb. Within three days, three interlocking realities became apparent that led to surrender and peace: first, the bomb existed and could be used again and again; second, the Russians would not be mediators but a new enemy; and third, the Americans would allow the institution of the Japanese emperor to be continued after surrender. No historian can tell us definitively how to weigh the impact of each of these three realities on the decision-making of the Japanese government, but there can be no doubt of their combined effect in producing quick surrender. It is also an open question how much this process shortened the war, but certainly there was a wholly genuine fear of heavy and sustained battlefield losses in the minds of US decision-makers. The more optimistic postwar estimates of some analysts do not tell us much. The broken Japan that they examined was not the one that US leaders or last-ditch Japanese generals had in

mind in July and early August 1945. At that climactic moment, the new bomb was seen in Washington as an enormously promising instrument of early victory over a ferociously determined enemy.

On the other hand, modern defenders

focus on a city demolished and a hundred thousand people indiscriminately killed denounce the deed and those who ordered it. That the war thus ended had itself already become that kind of war — as bad in its overall inhumanity as Hiroshima — is largely ignored on both sides. Yet the inescapable reality is that this kind of war was what had to be ended.

By this test, the bombings of Hiroshima on 6 August and Nagasaki on 9 August were a success. The Hiroshima bomb shocked the Russians into an immediate entry into the war, and these two events together shocked the Japanese government into a public offer of surrender on the single condition that the institution of the emperor be preserved. When it became clear that this condition was the one remaining obstacle to victory and peace, the reaction of US troops in the Pacific was decisive. The headline in the *New York Times* on 11 August said it all: "GI's in the Pacific Go Wild With Joy: 'Let 'em keep Emperor,' they say". The political doubts in Washington disappeared, and surrender terms were promptly agreed. As this history became known, it was recognized on all sides that the emperor himself played a decisive role in ending the war, deciding for peace when his ministers remained divided, and accepting the US condition that his own continued role must have democratic support. The emperor, more than any other one man, ended the war. But the situation in which he could become decisive was precipitated by the bomb on Hiroshima. The second bomb, on Nagasaki, strengthened the forces for surrender in the final Japanese debate. The bomb alone surely did not end the war, but just as surely it did precipitate the actions, Russian, Japanese and American, that together brought about surrender.

Later critics often neglect the reality that by the summer of 1945, the earliest possible ending of the war was a totally commanding purpose for Americans. If President Harry Truman had held back from any available means of rapid victory, he would have turned his back on his own strong convictions, abandoned the repeatedly proclaimed policy of his great predecessor, Franklin D. Roosevelt, and gone against the deeply held convictions of his countrymen, especially those in the armed forces still at risk. It is equally natural that millions of others—and especially the citizens of Hiroshima and Nagasaki—

Deliverance and destruction: the Hiroshima bomb can still sharply divide people.

of the decision to use the new weapon cannot usually defend the choice of large cities as targets. The best they can do is point out that the bombing of cities — in particular the use of deliberately massive incendiary attack — was already an established, accepted and highly destructive way of war before Hiroshima. The largest 'conventional' raid, on Tokyo in March 1945, was about equal to Hiroshima in its human cost, and such bombing of major cities continued throughout the spring and early summer of that year.

The actual course from nuclear decision to Japanese surrender has been largely neglected in this year's retrospective debate. A narrow but intensely felt disagreement did however erupt over the proposal to display Enola Gay, the plane that dropped the Hiroshima bomb, in a Smithsonian exhibition intended to commemorate the end of the war (see *Nature* 373, 371; 1995). Those who celebrate the bomber's mission and crew remember a great war swiftly ended by an extraordinary combination of nuclear technology with aircrew skill and courage. Those who

IMAGE
UNAMIGABLE
FOR A CRYING
FOR A CRYING
REASONS

Hulton Deutsch

remember most what happened to those two cities. These contrasting memories are reflected in the divided US sentiment on what to emphasize in anniversary commemorations of the end of the war, and the division is both understandable and honourable. (I should declare my own interest — I learned of Hiroshima when the infantry regiment of which I was a member was reaching the US West Coast on its way from Europe to Japan. To us it was wonderful news.)

One other decision deserves to be remembered. The target-selectors had a preferred target, a great industrial centre named Kyoto. But Henry Stimson, the secretary of war, knew that Kyoto was also the ancient capital of Japan; he had visited its monuments and seen its historic and artistic treasures, and he knew that to the Japanese it was a combination of Athens and Jerusalem. He struck it off the list, and with Truman's support he overruled appeals and blocked end-runs from General Leslie Groves. This great decision does not of itself excuse the bombing of other cities, but it does remind us that it is good to think hard about what we should not do with this weapon.

What is most remarkable about Hiroshima and Nagasaki is that they are the only times that nuclear weapons have so far been used in the world. Moreover, there is a worldwide consensus that this 50-year record of non-use is an achievement that should be extended. We have had sharp debates about ways and means, and we have had crises with nuclear danger in them, but the cumulative lesson of these debates and crises is that it has been right, for all concerned, to stay clear of the nuclear cliff. That we have managed to do so owes much to the sanity of statesmen, but something also to good luck, because there were dangers such as tactical bombs that we did not know about or take account of. Fortunately, we did not trigger them, but we might have — as in the Cuban missile crisis of 1962, a third of the way along those 50 years.

The Cuban missile crisis is probably the most serious moment of nuclear danger since Nagasaki. It sharply engaged the two states with the largest nuclear forces, and the encounter took place on an island where local Soviet commanders could have used tactical nuclear weapons. As the crisis intensified, the possibility of inflammatory conventional hostilities on the island became real, and with it the possibility of escalation completely unintended by the two heads of government. One may believe, as I do, that even if there had been a few local explosions, general war would have been avoided, but the legacy of anger and hatred would have been enormous. Fortunately, the general caution that governed both John F. Kennedy and Nikita Khrushchev was sufficient to prevent such an unintended

escalation. The escape was too narrow to leave either side with any appetite for additional encounters of this kind, and that is the lesson both sides learned. After Cuba, there would be no further US-Soviet nuclear crisis.

Still the hardest test of all for Americans may have come later, in the one large war the United States ever clearly lost: Vietnam. The cruel frustrations of the Vietnam years are still vivid in the memories of most adult Americans. But what is often forgotten about that war is that in all its long trials, only an occasional aberrant voice was raised in favour of a nuclear solution. Far over on the right, Senator Barry Goldwater once or twice sounded that way, but when he did he lost ground

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS REASONS

Enola Gay: those who celebrate the bomber's mission and crew remember an appalling war swiftly ended.

politically and shocked his own advisers. In 1967, the joint chiefs of staff strongly pressed in Congress for an increase in the bombing of North Vietnam, but they limited their appeal to conventional weapons. Their supporters in the Senate were careful not to press them further. As a real political choice the use of nuclear weapons, and even threats of nuclear weapons, remained off limits.

President Richard Nixon believed in nuclear threats — he thought, probably wrongly, that in 1953, as vice-president, he had seen them used effectively by President Dwight Eisenhower over Korea. But what was most important about Nixon's own inclination toward nuclear threat over Vietnam is that he could not act on it; he kept his warnings vague, and the contrary mindset of the American people was as plain to Hanoi as to him. Vietnam offers many painful lessons in what is not sensible in trying to help weak friends, and all sorts of alternative lines of policy were urged along the way. I know of no strong public argument, made at the time or later, that it would have been better to use nuclear weapons. Indeed, the domi-

nant nuclear lesson for US military planners from the experience has been that if they could not use nuclear weapons in Vietnam, they had better not count on using them anywhere, unless of course some one should use them on the United States first.

The first 50 years show many other crises with a nuclear thread in them, on matters as varied as the control of islands off China and border disputes in the Persian Gulf. What is striking as one looks back at such crises is that although their specific histories are varied, so that students can draw their own lessons about the importance of firmness or flexibility, there does seem always to have been a double awareness: that this situation could in fact turn nuclear, and that any such result would be a disaster for all concerned. My own impression — it is not a subject that lends itself to precision — is that the restraining power of this awareness has been great.

The end of the Cold War made possible — even easy — a dramatic reduction in the intensity of the nuclear competition between Moscow and Washington. At the end of the 1980s, under the leadership of Mikhail Gorbachev, the Russians quite simply abandoned the Cold War. Most conspicuously, they reversed their policy toward eastern Europe, where that war had begun, and the 40-year confrontation in the centre of Europe came to an end. Nuclear *détente* was not far behind.

The first consequence, as rapid as it was unexpected, was a shared recognition that the enormous nuclear armament of the superpowers made no sense for either of them. The change that began with Gorbachev got its most dramatic exposition from President Ronald Reagan. While the Cold War lasted, Reagan had been a partisan of nuclear strength, offensive and defensive. In particular he had supported a wildly unrealistic commitment to strategic defence, going so far as to say that such defences would protect the United States in the same way that a roof keeps rain out of one's house. In the new political atmosphere of the Gorbachev era, Reagan reversed his policy and turned against the nuclear arms race. Recognizing that neither of the two nuclear superpowers had a prospect of achieving any decisive superiority and that any large-scale nuclear exchange between them would be a shared catastrophe beyond all human experience, he pronounced in 1985 that "a nuclear war cannot be won and must never be fought". For the rest of his time in office, repeating this statement at frequent intervals, he joined with Gorbachev not only in large-scale agreements for nuclear arms reduction but also in dramatic political demonstration of a new trust between Moscow and Washington. Today there is still unfinished work to be

Hulton Deutsch

done on lowering the nuclear danger from superpower stockpiles, much of it interconnected with the political disintegration of the Soviet Union.

The meaning of the Reagan dictum goes far beyond the old competition between Washington and Moscow. That competition was the largest, the most wasteful and politically the most productive of fear and mistrust, but the Reagan proposition is valid for much smaller forces too. If two countries with small arsenals should use even five or ten nuclear weapons against each other, they would both be enormous losers. Exchanges of nuclear warheads are inevitable catastrophes. The choice of nuclear war, when both sides have survivable weapons, is an act of suicidal folly.

What is astonishing in retrospect is that political recognition of this proposition, and acts of arms control consistent with it, were so long in coming. There were many who understood the hazard from the beginning, but for some 40 years they did not prevail in either country. There was too much strength in the interlocking influence of traditional military thought and intense political competition. 'Their' weapons are bad and dangerous because 'they' are bad and dangerous, and so 'we' must keep 'ahead', offensively or defensively or both. The notion that enough can be enough, even if a rival may have more, does not have a large place in the inherited military thought of either Americans or Russians. The Cold War nuclear competition was predestined by the strategic predispositions of the two superpowers, and its peaceful reversal may be the most surprising thing about it.

Restraint

When we look more broadly at the overall record of all the nuclear-weapons states, we find a similar record of restraint. None of these states has used a warhead in anger, and no third state has come close to the enormous over-arming of the United States or the Soviet Union. None has sought nuclear confrontation with even a non-nuclear opponent. In the main, the available evidence supports the conclusion that the dominant purpose behind the nuclear forces so far deployed in all countries is defensive — that it is mainly fear that has been at work. This does not allow us to ascribe only defensive motives to all possible proliferators, but it does argue that there can be political ways and means — not unlike those used by Reagan and Gorbachev — to lower the level of nuclear competition region by region around the world. That has already happened in South America, where Argentina and Brazil have turned away from competition in nuclear-weapons development. Even in South Asia, where there is a tangled interconnection of nuclear-weapons programmes among Pakistan,

India and China, restraint is more remarkable than deployment, and fear more powerful than ambition.

An important lesson from lesser Cold War crises is that when a situation is still one of limited political commitment, it is highly unlikely that nuclear weapons will be heavily engaged. In Korea in 1950–53, the political judgement that the war must be limited — that General Douglas MacArthur must not be allowed to spread it to China — itself made plain that this war did not justify a resort to nuclear weapons. The United States could hold its ground without such weapons, and it did not want to run higher risks for larger stakes. In the same case, Stalin was already limiting himself still further, to a posture of "lets you and him fight". In Cuba, the crisis itself was nuclear, but that only made it more important that it be resolved by conventional strength; the obvious danger of nuclear confrontation was a powerful pressure for moderation on both sides. It is not wise to make a general law out of particular examples, each with its own peculiarities, but it remains a reasonable proposition that in a crisis over issues less important than national survival, the danger inherent in escalation will tend to keep conflict below the nuclear level.

There is continuing debate among specialists about the wisdom of recommending to all nuclear states a firm public policy that they will never be the first to use nuclear weapons, reserving them only for retaliation, and so deterrence of nuclear attack. Such a doctrine has the obvious advantage that if generally adopted and obeyed, it prevents nuclear war. Its disadvantage is that it does not meet real and deep-seated fears of conventional aggression. Such fears can be the main justification for a nuclear force. That was the situation of the North Atlantic Treaty Organization through most of the Cold War; it is the situation of Israel and Pakistan today. Moreover, some careful and moderate students believe that fear of what could happen *in extremis* has a value for peace-keeping that should not be weakened by a blanket pledge against first use. There are also careful and moderate commanders who would never say never to any choice that might one day become justified. I have been on both sides of this debate, which gets very narrow at its edge. I am content here to remark that there is an apparent rough justice in firing back that is absent when one fires first — so that any first use of nuclear weapons must have extraordinarily strong justification.

A broad generalization from the first 50 years is that openness is generally helpful to the moderation of nuclear competition. Any given nuclear capability will be more disturbing as an undiscussed but imperfect 'secret' than as an open

deployment. When forces are known with reasonable accuracy — and our capacity to have public knowledge about worldwide nuclear capabilities has never been higher than it is now — unreasonable fears are lowered. In particular, it is today possible for the five declared nuclear powers to know with assurance that none of them is in danger of having its ability to strike back knocked out by surprise attack. That knowledge gives justified confidence in Reagan's first principle of nuclear reality. The understanding of this proposition, along with a worldwide tradition of non-use unbroken since Nagasaki, is a good send-off for the second 50 years.

Clean-up

The clean-up of the warheads left over from the Soviet break-up will be costly, and sustained US help, along with parallel US reduction, is a good investment in the lessening of nuclear danger. The Nunn-Lugar Amendment of 1992 provided just such help for the post-Soviet nuclear clean-up; it deserves reinforcement, not abandonment, over the rest of the century.

This new phase will require hard work. This year has brought the welcome renewal of the Nuclear Non-Proliferation Treaty, but it was essentially a defensive effort, and it took a lot of effort not to lose ground. The agenda that needs action by many nations together is now long; we need much more than the mere survival of a valuable but imperfect instrument. But that is another subject, and the most that can be said here is that in the work of all the nations to limit proliferation, one essential requirement is the moderation and openness of the nuclear-weapon states. The renewal process created an appearance that the nuclear and non-nuclear states are opponents, whereas in reality they are partners who have a common enemy in the bomb itself. That indeed is the overriding lesson of Hiroshima, Nagasaki and the first 50 years. There is hard work ahead for every affected government, and all will be helped by remembering that whatever one may think of the way the bomb was used at the beginning, one cannot but give thanks that it has not been used since. Prudent men and women of all sorts, and especially those in authority, will do well to keep that history in mind as they continue to live with the possibility of nuclear catastrophe. The record of non-use that has been kept since Nagasaki is well worth handing on intact. It belongs to us all. □

*McGeorge Bundy is at the Carnegie Corporation of New York, 437 Madison Avenue, New York, New York 10022, USA. He is the author of *Danger and Survival* (1988), a detailed study of the nuclear age.*

Is the *Principia* publishable now?

Newton's theory of gravitation was a landmark in science but might easily have fallen foul of modern referees demanding explanations.

NEWTON'S theory of gravitation is one of the intellectual wonders of the world. Everybody knows the tale. Kepler's careful brooding over observations of the planets, notably those of Tycho Brahe, had led him to various generalizations about the characteristics of planetary orbits. Seeking the cause of these regularities, Newton came to the gravitational explanation — that each planet is bound to the Sun by a gravitational force inversely proportional to the square of the distance between them. The constant of proportionality is the product of the masses of the Sun and the planet with the universal constant G .

Because the dynamical response of a planet to the gravitational force is inversely proportional to its mass, the shape of a planetary orbit is independent of the mass of the planet, depending only on the object's specific angular momentum (angular momentum per unit mass). Newton went on to invent a rudimentary form of the differential calculus to demonstrate that closed planetary orbits are ellipses. He and others quickly used the gravitational law to deal with the Moon, and to show that spinning bodies such as the Earth should be flattened at the poles.

This recapitulation of the well known is partly stimulated by the appearance of the latest book by Subramaniam Chandrasekhar, the Nobel prize-winning theoretical astrophysicist and long-term editor of the *Astrophysical Journal*, who has worked through Newton's *Principia*, translating the great man's mathematics into modern form (see David Hughes' review on page 395 of this issue).

With the benefit of huge hindsight, we naturally regard the *Principia* as a landmark in the development of science, which it was. Even the controversy between the supporters of Newton and Leibnitz over proper attribution of priority for the invention of the calculus has long since been forgotten; the records show that the two men had reached the same result independently. And it was Newton, in any case, who for the first time applied the calculus to a problem that really mattered. The intuition that led him to the notion that an inverse square law would make ellipses of planetary orbits is nevertheless beyond ordinary ken.

Yet there is a flaw in the gravitational law at which Christiaan Huyghens, Newton's contemporary and intellectual equal, among others protested. What is all this about action at a distance? After all, Newton's law assumed that the gravitational

force was transmitted between a planet and the Sun without the intervention of any medium. By the same test, the force was transmitted instantaneously; increase the distance between a planet and the Sun and the attractive force is instantly decreased.

This aspect of the law of gravitation is plainly unphysical in the modern sense, but so it was before Newton's time. Had not the Ancients (Aristotle included) shrunk from the idea that a projectile could continue to move through the air without the continued intervention of some external influence? That is why people then invoked the idea that the movement of the projectile created a vacuum into which air would rush, impelling the projectile forward in its track. But Newton's first law deals magisterially with that issue: "a body will continue to move...".

Huyghens, being a Cartesian, had no doubt where matters stood on the law of gravitation. The idea of action at a distance could only be a kind of magic. Forces were transmitted from one object to another only on contact. So what would have happened if, in due course, the Royal Society had sent the manuscript of the *Principia* out to referees in the modern manner? The chances are that the exercise would have recruited a good many quizzical referees' reports. Anybody can write the dialogue; "by what means, pray, does the author fancy that this magic can be contrived over the great distance between the Sun and Jupiter and without the lapse of time?"

The other issue on which Newton and Huyghens were at odds is not irrelevant. Newton's corpuscular theory of light, the notion that light consists of "photons", was already public knowledge when the first edition of the *Principia* appeared. Indeed, Newton's contemporary fame rested more directly on his recovery of all the colours of the rainbow from a single beam of white light; that cannot have failed to enhance the reputation of the corpuscular theory.

Famously, Huyghens argued for the wave theory of light propagation. His demonstration of how a propagating wavefront can be constructed from wavelets emitted from the surface at an earlier time has a thoroughly modern ring. There are relics of it in most accounts of wave dispersion. But Huyghens was also vindicated (late in the eighteenth century) by the recognition that his expression for the index of refraction on going from one medium to another was, unlike Newton's, correct.

Certainly by the time Maxwell set about

putting electrodynamics on a mathematical basis in the 1860s, there was no doubt in his mind that action at a distance had to be avoided like the plague. In a paper published at the end of 1864, Maxwell referred to the earlier attempts by Wilhelm Weber to construct a mathematical theory of electromagnetism based on the assumption that there are particles with electrical and magnetic properties "which have the property of acting on one another at a distance by attraction or repulsion".

Dryly, Maxwell notes that Weber had been forced to suppose that the electrostatic force between two charges would depend on their relative velocity; he then declares "I have therefore preferred to seek an explanation . . . in another direction". That direction was of course the electromagnetic field which Maxwell took to be the source of electrical and magnetic forces. Following the familiar argument due to Huyghens, he concluded that there is an "aetherial medium filling space and permeating bodies and capable of being set in motion and of transmitting that motion from one part to another". That of course was the "luminiferous aether", the search for which occupied many of Maxwell's successors for much of the nineteenth century. In the end, the luminiferous aether collapsed under the weight of its internal contradictions.

Now, of course, Newton's corpuscular theory of light enjoys as much standing as Huyghens' account of wave motion in empty space. They are complementary aspects of light in the sense of quantum mechanics. It is nevertheless sobering that the sharp contrast of style between Newton and Huyghens has only now been resolved by the redefinition, in the language of quantum mechanics, of what the vacuum is.

Empty space is, in principle, a potential source both of photons and of pairs of material particles, electrons for example. Of course, there are strict rules to ensure that neither energy nor momentum appears without good reason, but at least for electromagnetic phenomena, a vacuum is a better version of the luminiferous aether than anybody at the end of the nineteenth century ever hoped to find. Meanwhile, it is intriguing but also important that the sharp conflict over action at a distance three centuries ago is one whose resolution depends crucially on quantum mechanics. Is it at all possible that a perceptive analysis of what, after all, was an honest conflict would have arrived at the truth about the nature of the vacuum much earlier?

John Maddox

Around and around we go

H. J. Melosh

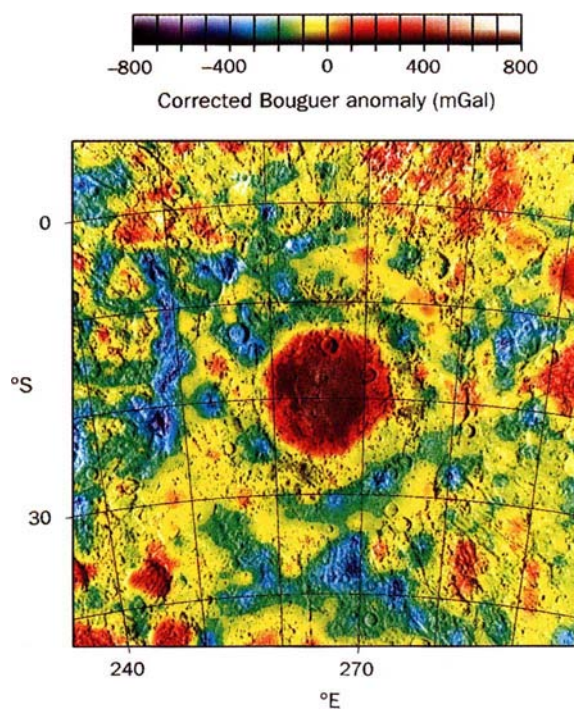
WAS the impact of an asteroid or comet big enough to gouge a 180-km-diameter crater into the Earth also big enough to wipe out the dinosaurs? In the wake of last summer's spectacular show when fragments of comet Shoemaker-Levy 9 hit Jupiter, many people would now answer with a resounding 'yes'.

However, a year before these impacts, the news media went wild¹ over the suggestion by Sharpton *et al.*² that the Chicxulub crater located in the Yucatán, now associated with extinctions at the Cretaceous-Tertiary (K/T) boundary, might be twice as big as previously thought. Because the impact energy scales as something more than the cube of the crater size, this would mean the knock-out punch was ten times more energetic than had been thought. But this larger crater size has been hotly contested ever since, engendering heated exchanges between Sharpton and others at scientific meetings. Now, Hildebrand *et al.*, using essentially the same data set, argue that the original diameter estimate of 180 km is indeed correct (page 415 of this issue³).

It may seem strange that experts cannot agree on something so basic as the size of an impact crater. A glance at the craters on the Moon makes it seem easy to measure their diameters, in that the freshest of them show sharp mountainous rims surrounding a central sunken plain. Nothing surely is simpler than to measure the distance between these rims, perhaps accounting for the inevitable irregularities of the rim crest, and present a number for the 'diameter'. To a large extent this is true, and there are now many tables of the 'rim crest diameter' of fresh craters on many planets and satellites (some cratering aficionados eschew this definition of diameter for the 'true diameter', which is the distance between the inner crater walls at the depth of the surrounding plain, but this is basically the same kind of measurement, only somewhat harder to make). The problem is that the structure of all craters is not the same and the relation of the final crater form to the impact that created it is often unclear.

It has been known for more than a century⁴ that the form of lunar craters changes as their size increases. Small craters exhibit a clean, bowl-shaped interior and sharp rim crest. Beneath their

floors lies a lens-shaped accumulation of broken rock. Larger craters (more than 15 km in diameter on the Moon, and from 2 to 4 km on Earth) have central peaks and are surrounded by a wreath of terraces marking the headscarps of great blocks of rock which slid into the crater cavity shortly after it formed. Their floors are covered with a layer of impact-melt rock which overlies a bed of broken rocks. In still larger craters the central peak is replaced by an interior ring of mountains.



Bouguer gravity over the Orientale basin on the Moon. The gravity anomalies (see colour scale) are overlaid on an airbrush map of the Moon's surface features. The most prominent ring is 620 km in diameter and its inner scarp is 6 km high. The large positive gravity anomaly in the basin's centre is due to the combined effect of a mantle upwarp beneath the basin and the attraction of the basalts in a small central mare, neither of which seems to be present beneath the Chicxulub crater on the Earth. Although the outer rings can be detected in the gravity field, they are not prominent. A linear regional trend has been removed from the data. (Image courtesy G. A. Neumann, Johns Hopkins University, and F. G. Lemoine, Goddard Space Flight Center.)

In the largest craters of all on the Moon, the impact site is encircled by one or more inward-facing scarps which probably represent the traces of large normal faults that developed as the entire lithosphere slid towards the crater cavity. One of the best-preserved examples of this type is the lunar Orientale basin (see figure), which has five well-defined rings ranging in diameter from 320 to 1,300 km. Which one represents the 'diameter' of the structure?

For the purposes of comparing crater size to impact energy, however, none of the diameter measurements described above can be used directly. Most of the complexity in crater form is due to different degrees of collapse after an initial crater is created by the force of the impact. This first crater starts out as a hemisphere around the impact site which grows until its bottom reaches some ultimate limit determined by gravity. The diameter continues to increase until it, too, is halted by gravity. The resulting cavity (called the 'transient crater') has a depth:diameter ratio of about 1:3, is highly unstable and immediately collapses under gravity to produce the variety of observed crater

forms which depend on the relative sizes of gravitational acceleration and on target properties such as strength and lithosphere thickness. It is really the transient crater size that matters in deducing impact energy, and so the problem is to relate the various measures of crater diameter to the transient crater size (although even this link is not totally secure due to uncertainties in the scaling relations⁵ that relate transient crater size to energy, and to such unknown parameters as projectile velocity, angle of impact and so on).

Determining the transient diameter of the Chicxulub crater is still more challenging because it is completely buried by up to a kilometre of later sediments, and must be studied by geophysical methods or by drilling. A fairly dense grid of more than 7,000 gravity stations is available on the northern Yucatán peninsula, and the crater itself shows up clearly as a circular, low Bouguer gravity anomaly almost 200 km in diameter, surrounding an interior central high. The low gravity is attributed to underlying low-density broken rocks, while the high gravity may represent denser impact-melted rock⁶.

Sharpton *et al.*² pointed to the existence of two ring-shaped gravity anomalies within the 200-km main basin with diameters of 104 and 154 km, one of which may correspond to the peak ring seen

in large lunar craters. In addition — and this is the main cause of the controversy — they identified a fourth ring with a diameter of 280 km, based on a number of small, isolated gravity highs outside the main low. They then argued that this outer ring is the true edge of the crater and inferred a transient crater size of 170 ± 25 km — because, they claimed, the steep gravity gradients at the edge of the main low could not be created by the conventional terrace zone and interior fill.

Hildebrand *et al.*³ use the same data set, supplemented by five closely spaced gravity profiles that they measured themselves. Their innovation is to plot the horizontal derivative of the Bouguer gravity instead of the gravity anomaly itself. This derivative removes regional trends and tends to highlight shallow structures, thus emphasizing anomalies due to the crater itself instead of the deeper features in the basement rocks. The crater shows up clearly in this gravity gradient plot, with a prominent ring of 180-km diameter (not substantially different from the 200-km diameter estimate above) and with several interior rings. But there is no hint of any larger ring. Hildebrand and colleagues are able to match the observed gravity profiles with a plausible density model in which the 180-km-diameter ring corresponds to the rim crest of a slumped crater (with perhaps an interior peak ring). In this case the transient crater was 80–90 km in diameter.

Does this new analysis completely settle the case for a 180-km-diameter crater? Alas, probably not. As the Clementine data⁷ show (see figure), the exterior rings of even such an indubitable ringed basin as Orientale do not show up clearly in gravity, a fact that Sharpton *et al.* should be quick to point out. On the other hand, Pilkington *et al.* note that two drill holes that nominally straddle the 280-km ring of Sharpton *et al.* encounter ejecta deposits at the same depth, instead of revealing the kilometre-high scarp that a ring would produce. The resolution of the issue of whether a 280-km diameter ring exists or not may require further drilling and seismic profiling, coupled with good old-fashioned stratigraphic correlation.

But in the end, the energy and lethality of the impact may not depend on whether the 280-km-diameter ring exists or not. For if this ring corresponds to the outer ring of a structure such as Orientale, it formed outside the rim of the crater and its diameter should not be used to infer the size of the transient crater. We are then back to the old problem of knowing just which diameter to use among the several that can be measured.

At present, it seems best to fall back upon the approach first proposed by Alvarez *et al.*⁸, in which the size of the projectile is deduced, not from the crater size (which had not been discovered at the

time), but from the quantity of extra-terrestrial iridium that was scattered over the Earth's surface. Since the estimated fluence of iridium at the K/T boundary has not changed substantially, neither should the estimate of a roughly 10-km-diameter projectile change. Such a projectile would produce a transient crater about 72 km in diameter, by the revised Schmidt–Holsapple scaling relation⁵, far more consistent with Hildebrand *et al.* than with Sharpton *et al.*

Was the impact of a 10-km-diameter comet or asteroid at 20 km s⁻¹ (or more) enough to cause the K/T extinctions? Many opinions have been changed by the impact of comet Shoemaker–Levy 9 on Jupiter, especially as it now seems that the largest individual fragments were only

about 700 m in diameter⁹. Although the impact velocity on Jupiter was 60 km s⁻¹, the difference in projectile size means that the fragments were at least 100 times less energetic than the K/T impact. The spectacular damage that they nonetheless caused underlines the tremendous destructive power of large impacts, a destructive power that is just now being analysed in detail. The current squabble over the detailed structure of Chicxulub is important to crater modellers, but no longer seems so relevant to the broader question of whether large impacts can cause biological extinctions. □

H. J. Melosh is in the Lunar and Planetary Sciences Laboratory, University of Arizona, Tucson, Arizona 85721, USA.

PULSAR ASTROPHYSICS

May or December?

Charles D. Bailyn

How old are binary and millisecond pulsars? Of the various issues relating to the origin and evolution of these enigmatic objects, this had seemed to be relatively uncontroversial. Most often, pulsar ages are derived by comparing their currently observed spin periods to their spin-down rates. For binary pulsars in which a companion white dwarf has been observed, the cooling timescale for the white dwarf can also be used to estimate the age of the system. Now, however, a binary pulsar has been discovered¹ in which these two age measurements are dramatically different, as Lorimer *et al.* describe on page 393 of this issue². The result is likely to add fresh uncertainty to an already confused situation.

Millisecond and binary pulsars are generally thought to be 'recycled' — once-dormant neutron stars which have been spun-up by accretion from a companion star. If that companion was a red giant star with a degenerate helium core, the core should remain after the accretion process has been completed. The end result would be a 'low-mass binary pulsar', a system containing a pulsar and a white-dwarf companion in a circular orbit with a period of a few hundred days.

The accretion stage should leave the pulsar on the so-called 'spin-up line', a theoretically predicted relation between the pulsar's magnetic field strength and the pulse period. After the accretion stops, the pulsar should slowly spin down, at a rate also determined by its magnetic field. From the present pulse period and spin-down rate, both of which can be measured to high accuracy, one can then infer the age of the pulsar since accretion stopped.

But the white dwarf is also evolving.

Such degenerate stars shine solely because of their residual internal energy — no fusion takes place to resupply energy to the star's core. Thus a white dwarf will cool as it shines at an easily calculable rate. Indeed, the predictions of simple models³ of the cooling rate of white dwarfs compare very impressively to modern observational data⁴. By determining the current temperature of a white dwarf, one can thus infer its age.

Starting in 1986, several white dwarfs in low-mass binary pulsars have been identified optically^{5–8}. Their temperatures were determined by multi-colour photometry, and the resulting cooling timescales were generally within a factor of a few of the derived spin-down ages of their companion pulsars. This was seen as a confirmation of the large ages (>10⁸ years) suggested by the pulsar spin-down rates, which in turn was taken as strong evidence that magnetic fields in pulsars do not decay below 10⁸–10⁹ gauss on timescales of about 10⁷ years, as had previously been suspected⁵.

Now Lorimer *et al.*² report similar observations of the white-dwarf companion of the recently discovered pulsar PSR J1012+5307 (ref. 1). The identification of the optical object as the white-dwarf companion of the pulsar seems secure: the positions are identical to high precision, and the distance inferred from the pulsar's dispersion measure is the same as that derived by comparing the white dwarf's luminosity to its observed flux. But in this case the white-dwarf cooling timescale of 300 million years is over an order of magnitude shorter than the pulsar's spin-down age.

What can have happened? The identification seems secure. Within the context

- Kerr, R. A. *Science* **261**, 1518–1519 (1993).
- Sharpton, V. L. *et al.* *Science* **261**, 1564–1567 (1993).
- Hildebrand, A. R., Pilkington, M., Connors, M., Ortiz-Aleman, C. & Chavez, R. E. *Nature* **376**, 415–417 (1995).
- Gilbert, G. K. *Bull. phil. Soc. Wash.* **12**, 241–292 (1983).
- Melosh, H. J. *Impact Cratering: A Geologic Process* (Oxford Univ. Press, New York, 1989).
- Pilkington, M., Hildebrand, A. R. & Ortiz-Aleman, C. J. *geophys. Res.* **99**, 13147–13162 (1994).
- Zuber, M. T., Smith, D. E., Lemoine, F. G. & Neumann, G. A. *Science* **266**, 1839–1843 (1994).
- Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. *Science* **208**, 1095–1108 (1980).
- Leonard, P. J. T. *Nature* **375**, 358 (1995).

of standard evolutionary models, the pulsar can hardly have spun down for billions of years before the creation of the white dwarf. The white-dwarf cooling timescale is certainly accurate to within a factor of two. If the magnetic field strength, and hence the spin-down rate, has recently dropped by a large factor, the currently inferred spin-down age might be anomalously long, but there does not seem to be any good reason to think that this is the case.

Possibly, however, the pulsar did not begin on the spin-up line, but instead near its current spin period of 5 milliseconds. This would imply that the age suggested by the white dwarf is in fact the system's true age. If many pulsars are younger than their spin-down ages, the pulsar birthrate would have to be higher than is generally calculated. This would exacerbate the

already existing problem that the required birthrate of recycled pulsars is greater than that of their purported precursors, the accreting low-mass X-ray binaries.

All of this provides some support for an alternative theory for the origin of recycled pulsars, namely the accretion-induced collapse of a white dwarf. In this picture, the pulsar is created when a mass-accreting white dwarf accretes sufficient mass from a companion star to collapse into a neutron star. Such an origin would avoid a low-mass X-ray binary stage, and thus alleviate the birthrate problem, and would also not be expected to result in a pulsar which begins its life on the theoretical spin-up line.

It is unlikely that a resolution to the evolutionary problems presented by PSR J1012+5307 can come from studying this individual object. Fortunately, pulsar

surveys are continuing apace, and many new recycled pulsars will doubtless soon be observed. Each new pulsar will contribute to determining whether the various evolutionary schemes which have been proposed are possible, probable or necessary. □

Charles D. Bailly is in the Department of Astronomy, Yale University, New Haven, Connecticut 06520, USA.

1. Nicastro, L. *et al.* *Mon. Not. R. astr. Soc.* **273**, L68–L70 (1995).
2. Lorimer, D. R., Lyne, A. G., Festin, L. & Nicastro, L. *Nature* **376**, 393 (1995).
3. Mestel, L. *Mon. Not. R. astr. Soc.* **112**, 583–596 (1952).
4. Winger, D. E. *et al.* *Astrophys. J.* **315**, L77–L81 (1987).
5. Kulkarni, S. R. *Astrophys. J.* **306**, L85–L89 (1986).
6. Bailly, C. D. *Astrophys. J.* **411**, L83–L85 (1993).
7. Bell, J. F., Bailes, M. & Bessell, M. S. *Nature* **364**, 603–605 (1993).
8. Danziger, I. J., Backer, D. & Della Valle, M. *Astr. Astrophys.* **276**, 382–388 (1993).

PHOTOSYNTHESIS

Short-circuiting the Z-scheme

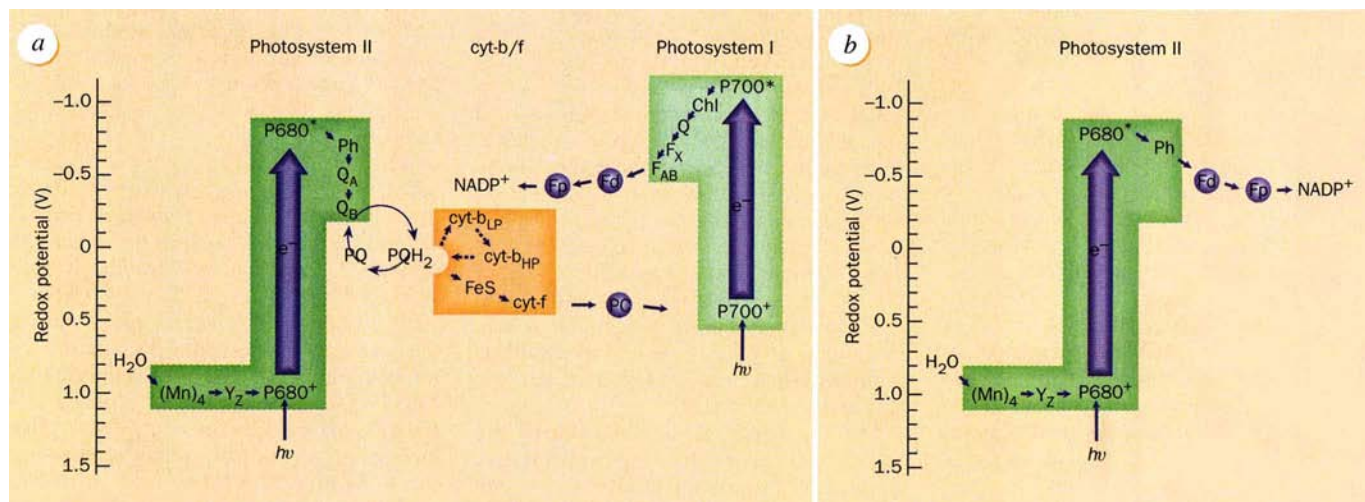
Jim Barber

PLANTS, algae and certain bacteria carry out the reactions of photosynthesis which are vital for sustaining life on our planet. The study of these reactions has followed a long and distinguished path, which reached a milestone with the introduction of the Z-scheme by Hill and Bendall¹ to explain how two cooperating photosystems, PSI and PSII, bring about electron transfer from water to carbon dioxide. Although a substantial body of experimental data exists to support this

concept, the work of Greenbaum *et al.*, reported on page 438 of this issue², suggests that under certain circumstances the fixation of carbon dioxide can be driven efficiently by PSII alone; moreover, the reaction mechanism involved could have significance for the evolution of aerobic photosynthesis.

The Z-scheme, depicted in *a* in the figure, is a well accepted concept, in which photosystem II, upon absorption of a photon, produces a sufficiently strong

oxidation potential to split water into molecular oxygen and reducing equivalents. In contrast, when photosystem I is excited it produces the low redox potential necessary to reduce NADP⁺ to NADPH. The two photosystems are embedded in the chloroplast membrane and are linked together by a chain of electron/proton carriers, including the mobile redox-active plastoquinone and plastocyanin, and the cofactors of the cytochrome *b/f* complex³. The light-driven charge separation within PSI and PSII, together with the interconnecting electron-transfer processes, generates a transmembrane electrochemical potential gradient of protons which powers the phosphorylation of ADP to ATP within an ATP-synthase



a, Electron flow pathway in oxygenic organisms presented in terms of redox potentials. Two different photosystems, I and II, cooperate to drive electrons from water to NADP⁺. This series of reactions is called the Z-scheme and was first proposed by Hill and Bendall 35 years ago¹. The cytochrome *b/f* complex acts as a redox link between PSII and PSI and uses a cyclic system to cope with the extra electron associated with the oxidation of PQH₂. PSI, PSII and cytochrome *b/f* are separate membrane-bound multiprotein complexes, represented here by three separate boxes². *b*, A modified scheme based on the results of Greenbaum *et al.* This shows how photosystem II alone can, under

certain conditions, directly reduce NADP⁺ and therefore bring about fixation of carbon dioxide. Abbreviations: Y_Z, redox-active tyrosine; P680 and P700, chlorophyll molecules that are the primary electron donors to PSII and PSI, respectively; Ph, pheophytin; Q_A and Q_B, bound plastoquinone molecules; PQ, mobile plastoquinone (PQH₂ is the reduced form of PQ); cyt, cytochrome — subscripts LP and HP indicate low and high potential, respectively; FeS, an iron–sulphur protein; PC, plastocyanin; Chl, chlorophyll; F_x, F_A and F_B are low-potential iron–sulphur centres; Fd, ferredoxin, Fp, flavoprotein (ferredoxin–NADP⁺ oxidoreductase); and NADP⁺, nicotinamide dinucleotide phosphate.

complex in the membrane. Together with the light-driven reduction of NADP^+ , this photophosphorylation reaction is required to convert carbon dioxide into organic molecules through a complex series of biosynthetic pathways.

It has often been argued that the co-operation of two distinct photosystems is necessary to bridge the large energy gap between water oxidation and NADP^+ reduction: this is likely to be in the region of 1.6 electron volts, given that several electron-transfer steps are needed in order to avoid reversibility. This is close to the 1.8 electron volts available for each photon absorbed by a single photosystem and is easily achieved by the combined energy of the two photosystems.

Despite this thermodynamic reasoning, D. I. Arnon, who with his co-workers discovered photophosphorylation⁴, challenged the validity of the Z-scheme⁵ for many years, arguing that water oxidation and NADP^+ reduction can be driven by PSII alone. The discovery that the first stable electron acceptor of PSII is a pheophytin molecule⁶, with a redox potential sufficient to reduce NADP^+ directly, added credibility to Arnon's proposals. Further support for his ideas emerged from experiments showing that PSII can directly bring about the reduction of NADP^+ under certain conditions⁷⁻¹⁰. Although the results substantiated the thermodynamic feasibility of this reaction, low quantum yields and the unusual nature of the experimental conditions combined to suggest that *in vivo* this alternative electron transfer pathway is unlikely to compete greatly with the normal Z-scheme reactions.

Greenbaum *et al.* have shown for the first time that the direct reduction of NADP^+ by PSII can occur at high rates and result in carbon dioxide fixation. They used a mutant of the green alga *Chlamydomonas reinhardtii* that contains PSII but has no PSI activity. Under anaerobic conditions, the mutant fixes carbon dioxide and grows photoautotrophically. NADP^+ reduction was detected in the mutant by the generation of hydrogen gas due to hydrogenase activity within the organism. These findings confirm that PSII alone can support photoautotrophic growth by generating the NADPH necessary for reducing carbon dioxide, vindicating the long-held views of Arnon (*b* in the figure). One surprise was that the quantum yields for NADP^+ reduction (as judged by hydrogen evolution) were about the same for both the mutant and the wild type.

The ability of the mutant depleted of PSI function to reduce NADP^+ , fix carbon dioxide and grow photoautotrophically disappeared when oxygen was present. It therefore seems that, under normal aerobic conditions, both PSII and PSI are required for photosynthesis: the

validity of the Z-scheme remains secure. But, as Greenbaum *et al.* point out, their results open up a discussion on how the water-splitting system of PSII evolved. The reaction centre of PSII shows remarkable structural and functional similarity to that of the purple photosynthetic bacteria, which grow anaerobically and so do not perform the water-splitting reaction. Could it be possible, argue Greenbaum *et al.*, that the evolution of the water-splitting process initially involved only a single photosystem — a modified version of the present-day purple-bacterial-type reaction centre — and that only later was PSI introduced, as the oxygen level in the atmosphere increased?

Their suggestion that the precursor to the Z-scheme was a PSII purple-bacterial-like water/ NADP^+ oxidoreductase system would have been much appreciated by Daniel Arnon, who died a matter of months before this paper was submitted for publication and to whom it is dedicated. It does not, however, detract from the well established dogma that under normal conditions the Z-scheme of Hill and Bendall is the preferred route of electron transport in oxygenic photo-

synthetic organisms. Moreover, the ideas of Greenbaum *et al.* on the evolution of the water-splitting system still need to be reconciled with the fact that the reaction centres of modern purple bacteria do not normally reduce NAD^+ directly, a reaction that is restricted to the PSI-like reaction centres of the anaerobic green-sulphur bacteria. □

Jim Barber is in the Department of Biochemistry, Imperial College of Science, Technology and Medicine, London SW7 2AY, UK.

- Hill, R. & Bendall, D. S. *Nature* **186**, 136–137 (1960).
- Greenbaum, E., Lee, J. W., Tevault, C. V., Blankinship, S. L. & Mets, L. J. *Nature* **376**, 438–441 (1995).
- Barber, J. & Andersson, B. *Nature* **370**, 31–34 (1994).
- Arnon, D. I., Allen, M. B. & Whatley, F. R. *Nature* **174**, 394–396 (1954).
- Arnon, D. I., Tsujimoto, H. & Tang, G. M. S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2942–2946 (1981).
- Klimov, V. V., Klevanik, A. V., Shuvalov, V. A. & Krasnovsky, A. A. *FEBS Lett.* **82**, 183–186 (1977).
- Klimov, V. V., Allakhverdiev, S. I. & Ladygin, V. G. *Photosyn. Res.* **10**, 355–361 (1986).
- Alsberson, P. A., Hsu, B.-D., Tang, G. M.-S. & Arnon, D. I. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3971–3975 (1983).
- Arnon, D. I. & Tang, G. M.-S. *FEBS Lett.* **253**, 253–256 (1989).
- Arnon, D. I. & Barber, J. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5930–5934 (1990).

EAST ANTARCTIC ICE SHEET

Stabilists strike again

Dick van der Wateren and Richard Hindmarsh

ON page 412 of this issue¹, Sugden and colleagues announce that glacier ice in the Dry valleys of East Antarctica seems to have survived for over 8 million years under a thin layer of sediment. This, they argue, testifies that stable polar conditions have persisted for at least that time. The data in the paper and the authors' claims will start the next round of a ten-year debate on the stability of the Antarctic ice sheet.

The Dry valleys are the largest ice-free area in Antarctica (see map), the proximity to McMurdo station and the abundance of rock outcrops making them the most intensely studied area in the entire continent. A number of geomorphological studies (see ref. 2 for review), supported by a wealth of radiometric ages, suggest that the landscapes of the Dry valleys region are very old. Surfaces, some of them quite steep, seem to have remained virtually unchanged for 15 million years. Their ages are constrained by ⁴⁰Ar/³⁹Ar dating of 50 presumed *in situ* vol-

canic ash deposits³ overlying these landforms. The Dry valleys represent a relict landscape, and erosion (presumably under a semi-arid climate) certainly occurred before the mid-Miocene². The implication is that Antarctic climate has been persistently cold since then, even during the warm mid-Pliocene. The preservation



View of Beacon valley with a lobe from the Taylor glacier pushing up the valley. The Miocene ice described by Sugden and colleagues¹ lies beneath the debris-covered floor of the valley.

Photo by G. Denton

of Miocene glacier ice is another strong argument in favour of the hypothesis that the East Antarctic ice sheet is stable.

The 'fossil ice' described by Sugden *et al.* is buried under a 50-cm-thick layer of sediment, interpreted as a glacial till derived from outside Beacon valley (the site studied; see photograph). The ash, dating of which put a minimum age on the till, was found in a sand-filled contraction crack in the till; it was supposedly deposited from the atmosphere into an open frost crack. Analyses of the ice fabric, its debris content and oxygen isotopes suggested that the ice is a remnant of an ancient outlet glacier of the East Antarctic ice sheet.

How thick would the ice have had to have been to survive to the present day, given that it will sublimate through the till? Sugden *et al.* compute a moisture flux assuming fully saturated conditions in the till, and find that about 1 m of ice would have sublimated in this time. A calculation based on a saturation gradient consistent with modern meteorological observations suggests that 1,000 m would have sublimated in this time — the implication, noteworthy in itself, is that the ice is rather younger. Another puzzling feature is that fabric survived in a stagnant ice mass. One would expect the ice to recrystallize slowly and the fabric to decay.

The long-running debate over Antarctic ice-sheet stability centres on the question of whether the East Antarctic ice sheet responded to global warming by maintaining more or less the same configuration, or whether it was substantially reduced in volume. The 'stable ice-sheet' hypothesis of George Denton and co-workers⁴ was challenged by the discovery, in the Beardmore glacier area and elsewhere in the Transantarctic Mountains (see map), of marine diatoms and *in situ* fossil terrestrial plants in thick sequences of glacial sediments at high elevation (the Sirius Group)^{5,6}. The marine diatoms span a number of biostratigraphical intervals, the youngest of which date from the Pliocene, a period of known global warming.

Together with evidence of palaeo-ice flow directions, these observations led to

the conclusion that the biostratigraphical intervals represent major deglaciations and marine incursions into what are currently subglacial basins in East Antarctica. A number of key diatoms, the ages of which were confirmed by a $^{40}\text{Ar}/^{39}\text{Ar}$ volcanic ash date of 3 million years from a drill core offshore from the Dry valleys⁷, constrained the latest of the deglaciations to 4–3 million years ago⁸. This 'dynamic

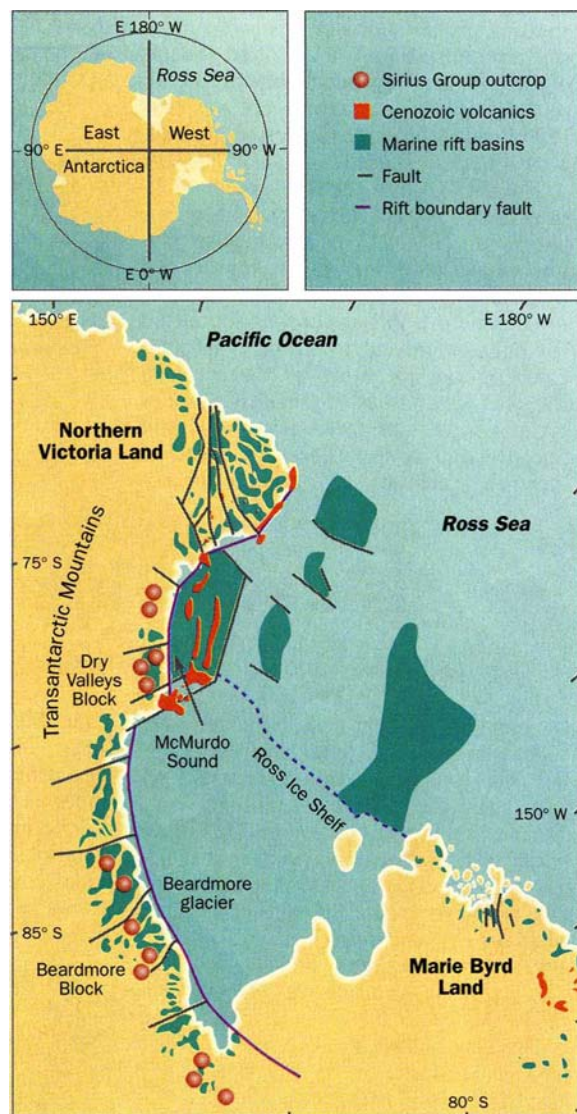
But the ice and overlying sediment (which they interpret as glacial till) could be remnants of rock glaciers, which are found about 1 km upvalley from the old-ice site. The till may have been reworked by the rock glacier and moved on top of the ice, which could then be of any age. Many reported ash sites^{3,9} are on slopes that are susceptible to soil creep, and some sand wedges seem to have been deflected downslope by creep. Other sites reported as "near *in situ*"⁹, such as avalanche cones and deposits on 30° slopes, can be interpreted in other ways.

The landscape of the Dry valleys has an evolutionary history which is quite different from that of other areas, and the new ash data presented by Sugden and members of his group^{1,3,9} may provide a solid chronostratigraphical background for studying that history. Part of this difference may be due to the tectonic history of the West Antarctic rift system, of which the Transantarctic Mountains are the uplifted western flank¹⁰. Differential surface uplift of fault blocks along the 3,000-km-long mountain range appears to lead to quite different denudation patterns and even produce variations in regional climate¹¹.

When geologists are faced with such severe problems of interpretation, they try to find a similar area where they can test their ideas. Unfortunately, the Dry valleys are unusual, and the problems of interpretation are so linked to the fundamental scientific question of whether it is a freeze-dried landscape from the Miocene, that such comparisons are impossible. Those bored with fieldwork in Antarctica will have already noted that the closest analogies may be on Mars. Alternatively, the offshore record is better preserved than the terrestrial record and has the advantage of a more direct correlation to both the deep ocean and the tectonic records.

For example, high-resolution seismic data from the McMurdo Sound (just offshore from the Dry valleys), and from the continental shelf of the western Ross Sea, reveal 20 seismic-stratigraphic sequences¹² representing separate grounding events or ice-sheet advances across the Ross Sea continental shelf. In McMurdo Sound ten of these are post mid-Miocene while in the eastern Ross Sea seven seismic sequences have been identified in the Plio-Pleistocene interval¹³. Continental shelf and margin records indicate highly dynamic behaviour of the East and West Antarctic ice sheets and are hard to reconcile with the 'stable ice-sheet' hypothesis, begging the question of how representative the Dry valleys' geomorphological record is for the climate history of Antarctica.

Issues such as these were discussed at the Pliocene Antarctic Glaciation Workshop, held at Woods Hole in April, the



Location map showing fault blocks and marine rift basins in the Cenozoic West Antarctic Rift system. Modified from ref. 11.

ice sheet' hypothesis of Webb and Harwood⁶ postulates a number of major fluctuations of an East Antarctic ice sheet with different meteorological and glaciological characteristics to the present ice sheet which developed in the late Pliocene. A major deglaciation of East Antarctica would lead to a rise in global sea level of several tens of metres.

The stabilist argument depends greatly on how the ashes were emplaced. Sugden and members of his group^{1,9} go to great lengths to show that the ashes are *in situ*.

object of which was to define new research strategies towards a solution of the stabilist–dynamist controversy. The use of quite different research methods — radiometric dating of geomorphology in one area and biostratigraphy in the other — has been a source of misunderstanding between the two groups. Confusion about sediment sampling, which for example leaves room for doubt over whether the Pliocene diatoms were transported by ice or wind, is another. A proposal to sample Sirius Group sediments at different depths below the surface to settle this matter finally was generally approved.

Although this and a previous meeting¹⁴ in 1992 greatly improved mutual understanding, the stabilist–dynamist controversy will not be resolved soon. Our knowledge is still severely limited by lack of geochronological data. The land record from Antarctica is quite fragmentary, seismic stratigraphies and biostratigraphies are in their early stages of development, and correlation of these records to lower latitudes and the Northern Hemisphere has yet to begin. Studies of the tectonic evolution of the uplifted rift flank and adjacent sedimentary basins^{10,15} should help, as should independent geochronologies and new exposure age-dating techniques. These need to be related to the geomorphological evolution of the flank, ice-sheet fluctuations, eustatic sea-level change and the ocean isotope record, while a deeper understanding of the influence of the East Antarctic ice sheet on global circulation is needed to relate these to the Pliocene in other continents. □

Dick van der Wateren is in the Research School of Sedimentary Geology, Institute for Earth Sciences, Free University, De Boelelaan 1085, NL 1081 HV Amsterdam, The Netherlands. Richard Hindmarsh is in the Ice and Climate Division, British Antarctic Survey, Madingley Road, Cambridge CB3 0ET, UK.

Ins and outs of the ribosome

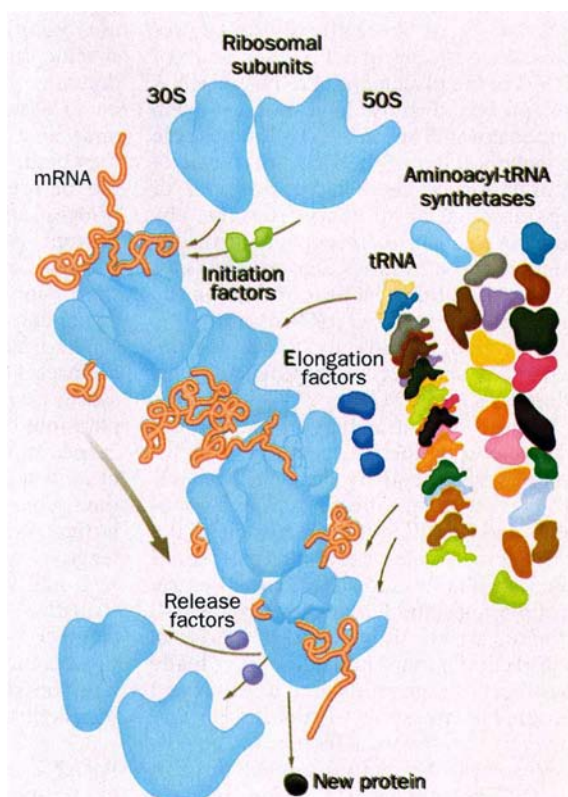
Robert A. Zimmermann

THE detailed morphology of ribosomes, long a matter of conjecture and controversy, has until recently eluded precise definition. These ribonucleoprotein particles are the cellular factories for protein biosynthesis, providing the sophisticated machinery that takes in the raw materials — amino acids linked to their transfer RNA carriers — and adds them one by one to growing polypeptide chains according to the genetic blueprints encoded in messenger RNA. Although the overall shape of *Escherichia coli* ribosomes has been known for some time from conventional electron microscopy, the limited resolution of this method prevents a proper appreciation of how the structure of the ribosome is adapted to its functional requirements.

On page 441 of this issue, Frank *et al.*¹ present a new view of *E. coli* ribosome structure based on a 25-Å reconstruction of several thousands of images obtained by energy-filtering cryo-electron microscopy. This new approach, involving the use of samples embedded in ice, preserves the ribosome in a hydrated form during the imaging process and permits the detection of subtle conformational features that would otherwise be lost. What is revealed is an astonishingly irregular structure replete with bulges, lobes, channels, bridges and even tunnels, many or all of which could be functionally significant. Satisfyingly, the overall conformation is a familiar one, with clear similarities to earlier visualization in the electron microscope^{2,3}.

To comprehend how this new information could help us understand the mechanism by which proteins are made, it is as well to recall the complexity of the *E. coli* ribosome and of the functions it performs. A total of 55 different proteins and three ribosomal RNAs, packed into 30S and 50S subunits in a manner that seems to lack any regular order, provide highly specific binding sites for the in-

coming aminoacyl-tRNA and the growing peptidyl-tRNA, as well as the site through which the discharged tRNA exits after it has completed its rounds. In addition, the ribosome must ensure that the mRNA remains bound during synthesis of the polypeptide chain and must mediate the transient interactions of the non-ribosomal protein factors that promote polypeptide chain initiation, elongation and termination (see figure). The main functions of the ribosome throughout this process are to guarantee the fidelity of interaction between the mRNA codons



Protein synthesis in *E. coli* involves over 125 macromolecules. Aminoacyl-tRNA synthetases link amino acids to their cognate tRNAs; the aminoacyl-tRNAs are then transported to the ribosomal A site in a complex with elongation factor Tu and GTP. Here the tRNA anticodon recognizes and binds to the complementary mRNA codon specifying the next amino acid in the growing polypeptide chain. The peptidyl moiety of peptidyl-tRNA in the ribosomal P site is joined to the A-site-bound aminoacyl-tRNA in a reaction catalysed by peptidyl transferase. The extended peptidyl-tRNA is translocated to the P site with the aid of elongation factor G. At the same time, the mRNA moves by three nucleotides to place a new codon in the A site. The discharged tRNA is transferred to the E (exit) site and dissociated from the ribosomal complex. Non-ribosomal initiation factors stimulate the start of polypeptide chain synthesis in response to special AUG codons, while non-ribosomal release factors promote termination of polypeptide chain growth in response to UAG, UAA and UGA termination codons, resulting in release of the completed protein and recycling of the ribosomal subunits. Figure modified from ref. 11.

1. Sugden, D. E. *et al.* *Nature* **376**, 412–414 (1995).
2. Denton, G. H., Sugden, D. E., Marchant, D. R., Hall, B. L. & Wilch, T. I. *Geogr. Annl. Stockh.* **75A**, 155–204 (1993).
3. Marchant, D. R., Denton, G. H., Swisher, C. C. & Potter, N. J. *Bull. geol. Soc. Am.* (in the press).
4. Denton, G. H., Bockheim, J. G., Wilson, S. C. & Stuiver, M. *Geology* **12**, 263–267 (1984).
5. Webb, P. N., Harwood, D. M., McKelvey, B. C., Mercer, J. H. & Stott, L. D. *Geology* **12**, 287–291 (1984).
6. Webb, P. N. & Harwood, D. M. *Quat. Sci. Rev.* **10**, 215–223 (1991).
7. Barrett, P. J., Adams, C. J., McIntosh, W. C., Swisher, C. C. & Wilson, G. S. *Nature* **359**, 816–818 (1992).
8. Wilson, G. S. *Quat. Sci. Rev.* **14**, 101–123 (1995).
9. Marchant, D. R., Denton, G. H. & Swisher, C. C. *Geogr. Annl. Stockh.* **75A**, 269–302 (1993).
10. van der Beek, P., Cloetingh, S. & Andriessen, P. *Earth planet. Sci. Lett.* **121**, 417–433 (1994).
11. van der Wateren, F. M., Luyendyk, B. P., Verbers, A. L. L. M. & Smith, C. H. *Terra Antarctica* **1**, 453–456 (1994).
12. Henrys, S. A., Bartek, L. R., Anderson, J. B. & Barrett, P. J. *Terra Antarctica* **1**, 373–374 (1994).
13. Shipp, S. *et al.* *Terra Antarctica* **1**, 381–384 (1994).
14. van der Wateren, F. M., Verbers, A. L. L. M. & Tessensohn, F. (eds) *LIRA Workshop on Landscape Evolution* (Rijks Geologische Dienst, Haarlem, 1994).
15. Cloetingh, S. & Kooi, H. *Terra Nova* **4**, 340–350 (1992).

which specify each of the amino acids and the complementary anticodon sequences of the appropriate aminoacyl-tRNAs at the 30S subunit decoding site; promote catalysis of the peptide bond between the resident peptidyl-tRNA and the incoming aminoacyl-tRNA by an activity located in the 50S subunit; and carry out the coordinated translocation of the mRNA and the newly extended peptidyl-tRNA from the aminoacyl-tRNA site to the peptidyl-tRNA site.

Frank *et al.*¹ have assigned many of these functional sites on their model by analogy with what is known from lower-resolution ribosome models, using characteristic morphological features as guides. These identifications must be regarded with some caution, however, as substantive landmarks on the new structure, such as the locations of specific ribosomal proteins, the positions of defined segments of rRNA or the placement of tRNA, mRNA and protein ligands, have not yet been pinned down. Therefore, it is not possible to make full use of the vast storehouse of information on the relative positions of the ribosomal constituents determined by immune electron microscopy, neutron diffraction and intramolecular crosslinks²⁻⁶. By the same token, a host of data on the proximity of mRNA, tRNA, antibiotic and protein ligands to these components⁶⁻⁸ cannot yet be employed to best advantage.

Despite the difficulties in making definitive correlations of the morphological features visualized by cryo-electron microscopy with specific ribosomal components and ligands at this time, some of the former, such as the head and platform of the 30S subunit, and the L1 arm, central protuberance and L7/L12 stalk of the 50S subunit, are so distinctive that there can be little doubt that Frank *et al.*¹ have made the correct assignments in these cases and reasonable proposals in others. Perhaps the most controversial feature of the new model is the bifurcating tunnel through the 50S subunit, which is imagined to serve as the exit route for nascent polypeptide chains. In particular, implication of this tunnel in protein export is problematical because, in contrast to the secretory process in eukaryotes, there is no compelling evidence that the *E. coli* ribosome docks with specific receptors in the cytoplasmic membrane⁹.

Interest in the prokaryotic ribosome has remained high, in part because of the tantalizing expectation that sustained and diligent effort will lead to an intimate understanding of the way in which this complex organelle mediates protein synthesis. The *E. coli* prototype clearly has the advantage here as no eukaryotic ribosome has been subjected to the same level of scrutiny. Almost every aspect of the *E. coli* ribosome has been examined, from the protein-rRNA interactions that

underlie assembly to the functional role of individual nucleotides in synthesis of the peptide bond, and it is from prokaryotes that inferences about the key role of rRNA in ribosome function have been drawn¹⁰. More generally, it has become apparent that the basics of ribosome structure and protein biosynthesis have been conserved to a remarkable extent throughout evolution. Thus, there is a strong belief that a detailed understanding of the *E. coli* ribosome should be of wide applicability to ribosomes from all organisms.

Although the path ahead is daunting, the detailed model described by Frank *et al.*¹ offers a promising foundation for future experimentation. The first step might be to determine how the morphological features seen in cryo-electron microscopy reconstructions relate to specific ribosomal proteins and specific domains of the rRNA. A second could be to ascertain the locations of the ligands and the functional sites to which they bind. A third would be to investigate the nature of the multiple intersubunit bridges and of the tunnel(s) that the nascent polypeptides are supposed to traverse while avoiding unfavourable steric contacts with surrounding ribosomal constituents.

When all is said and done, what Frank *et al.*¹ present is not really a model of protein synthesis in a mechanistic sense, but a plausible depiction of the apparatus that carries out protein synthesis. While the new structural insights are important in themselves, they draw attention to the intricacy of the ribosome and the challenge of making sense of this marvelous machine in all its complexity. The structure of the ribosome as derived from cryo-electron microscopy is likely to become the new pattern for modelling the ribosome and ribosome-ligand interactions. □

Robert A. Zimmermann is in the Department of Biochemistry and Molecular Biology, University of Massachusetts, Amherst, Massachusetts 01003, USA.

1. Frank, J. *et al.* *Nature* **376**, 441-444 (1995).
2. Stöffler-Meilicke, M. & Stöffler, G. in *The Ribosome: Structure, Function and Evolution* (eds Hill, W. E. *et al.*) 123-133 (Am. Soc. Microbiol., Washington DC, 1990).
3. Oakes, M. I. *et al.* in *The Ribosome: Structure, Function and Evolution* (eds Hill, W. E. *et al.*) 180-193 (Am. Soc. Microbiol., Washington DC, 1990).
4. Capel, M. S. *et al.* *Science* **238**, 1403-1406 (1987).
5. Walieczech, J. *et al.* *Biochemistry* **28**, 4099-4105 (1989).
6. Brimacombe, R. *Eur. J. Biochem.* **230**, 365-383 (1995).
7. Noller, H. F. *et al.* in *The Ribosome: Structure, Function and Evolution* (eds Hill, W. E. *et al.*) 73-92 (Am. Soc. Microbiol., Washington DC, 1990).
8. Wower, J. *et al.* in *The Translational Apparatus* (eds Nierhaus, K. H. *et al.*) 455-464 (Plenum, New York, 1993).
9. Wolin, S. L. *Cell* **77**, 787-790 (1994).
10. Noller, H. F. *et al.* *Rev. Biochem.* **60**, 191-227 (1991).
11. Goodsell, D. S. *The Machinery of Life* p. 47 (Springer, New York, 1993).

Star Wars 2

In the early days of radar, there were hopes that it might be a weapon as well as a detector. A strong radar beam, it was thought, might usefully sabotage the electronics of an enemy aircraft. This soon proved unfeasible. Daedalus is now reviving the idea.

A conventional radar beam diverges too fast to be an effective weapon. Even with the biggest steerable dish-aerial, 100 kW of power at the transmitter is diluted to 1 W per square metre or less at the target. But Daedalus recalls the aperture-synthesis technique used by radioastronomers. A network of aerials spread out over a large area, but coupled together as a phase-coherent receiver, has the angular resolution of a single aerial as big as the whole area containing the network. Modern synthetic-aperture systems can locate a radio-source to within a tiny fraction of an arcsecond. So Daedalus plans to run such an array in reverse, as a transmitter — which will, of course, have the same angular resolution. A set of radar transmitters distributed over a few thousand square kilometres, worked as a synthetic-aperture array and aimed to converge at a distant aircraft, could focus on it with awesome precision. An array resolving to 0.1 arcsec could reach out 100 km and still concentrate its power in a circle 10 cm across.

Focused to this intensity, 100 kW of microwave energy would melt a hole in any aircraft. Even a brief or glancing strike would wreck its avionics. The beam could be aimed and steered by instant computer-generated phase-shift commands. An aircraft could be destroyed as soon as sighted, and even a big fleet of them could be knocked out in rapid succession. A really big synthetic-aperture radar could even attack missiles and satellites in space, making a 'Star Wars' defence feasible at last.

As a weapon, this ultimate in phased arrays is fairly benign. It threatens nobody, and can only be used to defend the area over which it is distributed. Terrorists could not steal it or hijack it. Even if they took over one aerial installation they couldn't focus it without the others.

The system might even have peaceful uses. An electric model aircraft has already been powered by a microwave beam aimed at it from below. The technology could easily be scaled up. It should be feasible to fit a light aircraft with microwave-rectifying wings, and use their output to drive electric propellers. The resulting craft would need no fuel and would eject no exhaust. Simple, silent and safe, it could transform short-haul civil aviation.

David Jones

Birth rate of millisecond pulsars

SIR — The characteristic ages of millisecond pulsars are determined from their spin-down rates, but these might not be good estimates of their true ages. We recently reported¹ a millisecond binary pulsar (J1012+5307) and tentatively identified the companion (star X) as a white dwarf. Here we present optical colours and astrometry that confirm star X as the pulsar's companion, and that it is a white dwarf. The cooling age of the white dwarf is only ~300 million years, which is an order of magnitude less than the pulsar's characteristic age. If characteristic ages are generally in error by this amount, the birth rate of millisecond pulsars must be much greater than previously estimated.

The field containing PSR J1012+5307 and star X (see Fig. 2 of ref. 1) was

From the JKT and Schmidt data, using PPM guide-star positions, the position of star X at this epoch was found to be RA (2000.0) = 10 h 12 min 33.41 s, dec. (J2000.0) = 53° 07' 02.7". Further radio observations of PSR J1012+5307 with the 76-m Lovell Telescope¹ have resulted in significant improvements in the accuracy of the position of the pulsar, its period derivative and its orbital elements (see table). Together, these data show a position difference between PSR J1012+5307 and star X of only 0.2 ± 0.5 arcsec for epochs less than one year apart, making the identification almost certain, given the low stellar density in the field.

We obtain an upper limit of 33 mas yr⁻¹ for the proper motion of star X from this new position and that reported in ref. 1

from the Palomar Sky Survey taken almost 40 years earlier. This implies a transverse speed ≤ 75 km s⁻¹ at the distance, d , of 520 pc inferred from the dispersion measure of J1012+5307.

The pulsar shows strong interstellar scintillation, providing a rough estimate of the transverse speed of the binary system. Using standard analysis techniques⁴, we estimate this to

be ~50 km s⁻¹, consistent with the upper limit for star X. Because a transverse speed, v , gives rise to an overestimate of the true period derivative⁵ of $\Delta P = Pv^2/cd$, we find that a better estimate of the intrinsic characteristic age $\tau = P/2\dot{P}$ of PSR J1012+5307 is 7 Gyr.

The $B-V$, $V-R$ and $V-I$ colours of star X are all consistent with a black-body temperature of $9,400 \pm 300$ K. Adopting the mass-radius relation reported in ref. 6, we estimate the luminosity of a $0.15 M_{\odot}$ white dwarf at this temperature to be $L \approx 4.2 \times 10^{-3} L_{\odot}$. At a distance of 520 pc, this implies a visual magnitude of 19.3, in good agreement with the observed value. Thus both the colour and magnitude of star X are consistent with it being a 9,400-K white dwarf star.

Taking the luminosity $L_* = 4.2 \times 10^{-3} L_{\odot}$ from the cooling curve in ref. 7, we estimate the age of star X to be ~300 Myr — more than a factor of 20 below the characteristic age of J1012+5307. Given the strong evidence in favour of the association, and the relatively small separation of the binary pair (~0.3 $R_{\odot}$), the discrepancy between the observed and predicted luminosities might be due to the heating of star X by the spin-down energy loss of J1012+5307. However, a simple calculation shows that this makes only a minor contribution (~10⁻² L_*) to the stellar luminosity. The discrepancy is surely due to the

system being much younger than the pulsar characteristic age, implying that the initial spin period of J1012+5307 was similar to the presently observed value⁸. On much larger time-scales, similar, but less certain, discrepancies have been noted⁹⁻¹¹ for the millisecond pulsar J0437-4715 and the anomalously high characteristic ages of some Galactic-disk millisecond pulsars which exceed the age of the Galaxy⁸.

This identification underlines the possible substantial overestimate of the ages of millisecond pulsars by use of their characteristic ages. If this is generally true, the birth rate required to sustain the observed Galactic population will be much greater than previous calculations¹² have suggested.

D. R. Lorimer, A. G. Lyne

*The University of Manchester,
Nuffield Radio Astronomy Laboratories,
Jodrell Bank, Macclesfield SK11 9DL, UK*

L. Festin

*Astronomical Observatory,
Box 515, S-751 20, Sweden*

L. Nicastro

*Istituto di Tecnologie e Studii delle
Radiazioni Extraterrestri - CNR,
Via Gobetti 101, 40129 Bologna, Italy*

RADIO TIMING SOLUTION OF PSR J1012+5307

RA (J2000)	10 h 12 min 33.4326(4) s
Dec. (J2000)	53°07'02.66(1)''
Period	0.00525574901198(2) s
Dispersion measure	9.0207(4) cm ⁻³ pc
Observed period derivative	$1.46(8) \times 10^{-20}$
Orbital period	0.604672713(5) d
Orbit semi-major axis	0.581816(6) light s
Orbit eccentricity	$< 2 \times 10^{-5}$
Epoch of period and ascending node (MJD)	49220.447499(1)

Numbers in parentheses represent twice the standard errors in the last quoted digit, inferred from a formal fit to the data.

observed on 21 October 1994 and 1 January 1995 during staff time at the Nordic Optical Telescope (NOT), on 22 January 1995 during service time on the Jacobus Kapteyn Telescope (JKT) and on 3 February 1995 during staff time at the Schmidt telescope at the Kvistaberg observatory. With ~1 arcsec seeing, we obtained 300-s exposures in B, V, R and I filters on the NOT; 1,000 s in V, R and I on the JKT; and 3,600 s in R at the Schmidt telescope. We obtained photometry for star X from the NOT data by using the curve-of-growth technique² in the PHOT procedure within the IRAF facility. We found that $V = 19.58 \pm 0.02$, $B-V = 0.20 \pm 0.04$, $V-R = 0.09 \pm 0.03$ and $V-I = 0.26 \pm 0.03$. Given the galactic coordinates of J1012+5307 ($l = 160^\circ$, $b = 51^\circ$), any reddening effects on these data are negligible³.

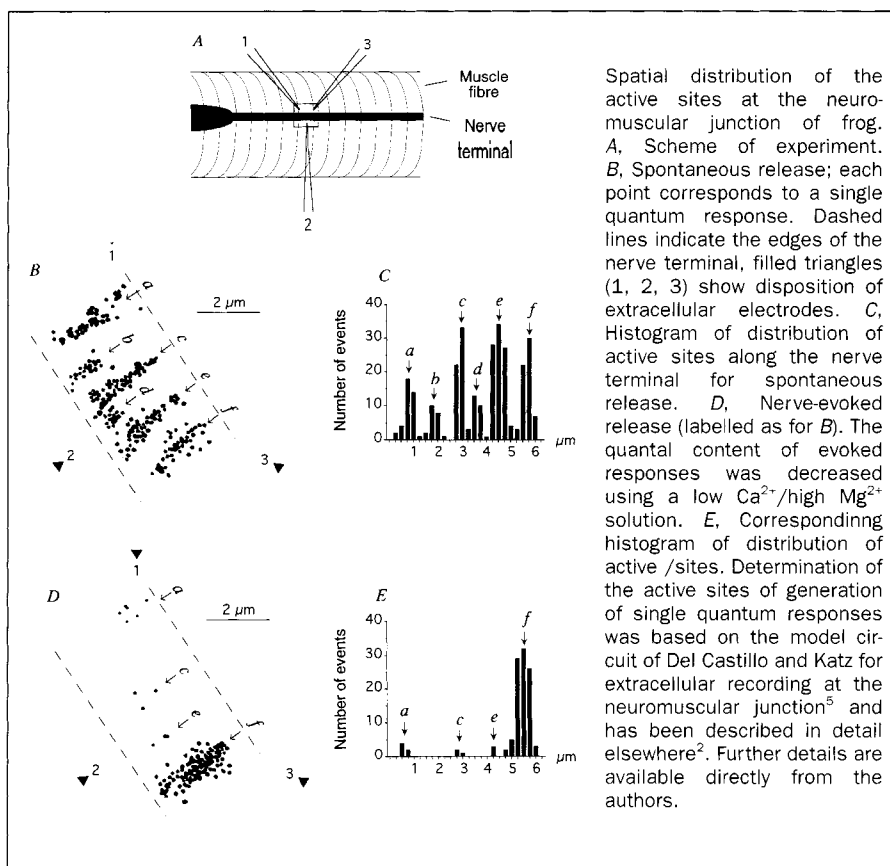
Localization of active zones

SIR — Synaptic transmission is believed to proceed by the function of spatially distinct zones, which can be clearly identified by morphological studies¹. This hypothesis, however, lacks direct functional evidence. Here we show that sites of generation of postsynaptic responses evoked by a single quantum of transmitter at the neuromuscular junction of frog are grouped in active zones with spatial parameters similar to those of morphologically determined active zones.

Extracellular miniature endplate potentials (e.m.e.p.ps) were simultaneously recorded using three independent electrodes placed perisynaptically (4 in the figure). In controls we found that the amplitude of e.m.e.p.ps was reciprocal to the distance between the electrode and the site of synaptic current generation (active site)²; thus, comparison of the amplitudes of unitary e.m.e.p.ps recorded by the three electrodes allowed us to determine the location of active sites in two dimensions.

Sites of spontaneous release were grouped in active zones dispersed across the nerve terminal (B in the figure; zones $a-f$). The probability of localization of the active sites along the nerve terminal also revealed peaks corresponding to the active zones (C in the figure). The distance between the active zones was 1.3 ± 0.5 mm while their length was 1.9 ± 0.9 mm ($n=17$), which corresponds well with those determined by morphological studies^{3,4}. Some of the active zones did not extend

1. Nicastro, L. *et al.* *Mon. Not. R. astr. Soc.* **273**, L68-L70 (1995).
2. Stetson, P. B. *Publs astr. Soc. Pacif.* **102**, 932-948 (1990).
3. Burstein, D. & Heiles, C. *Astr. J.* **87**, 1165-1189 (1982).
4. Lyne, A. G. & Smith, F. G. *Nature* **298**, 825-827 (1982).
5. Shklovskii, I. S. *Sov. Astr.* **15**, 562-565 (1970).
6. Savonije, G. J. *Nature* **304**, 422-423 (1983).
7. Iben, I. & Iutikov, A. V. *Astrophys. J.* **311**, 742-752 (1986).
8. Camilo, F., Thorsett, S. E. & Kulkarni, S. R. *Astrophys. J.* **421**, L15-L18 (1994).
9. Bell, J. F., Bailes, M. & Bessell, M. S. *Nature* **364**, 603-605 (1993).
10. Bailyn, C. D. *Astrophys. J.* **411**, L83-L85 (1993).
11. Danziger, I. J., Baade, D. & Della Valle, M. *Astr. Astrophys.* **276**, 382-388 (1993).
12. Lorimer, D. R. *et al.* *Astrophys. J.* **459**, 933-938 (1995).



across the entire width of the presynaptic terminal (*B* in the figure, zones *b* and *d*), and may correspond to morphologically identified, 'fragmented' active zones³. Densities of active sites were the same for all active zones, indicating that the level of spontaneous transmitter release is relatively uniform. The spatial distribution of the active sites of responses evoked by presynaptic action potentials had similar profiles to those spontaneously generated (*D*, *E* in the figure). Active sites of evoked release were grouped in active zones similar to those of spontaneous release and coincided with active zones of spontaneous release recorded in the same experiments ($n=7$; *B*, *C* in figure). In contrast to spontaneous release, the probability of evoked response generation was significantly different between active zones: in some it was high (*E* in the figure, zone *f*) whereas others were almost silent (zones *a-e*).

The present study provides direct functional evidence for the hypothesis that

synaptic transmission results from the function of spatially separated active zones. The approach described here has substantial advantages over other techniques⁵⁻⁷ and can be used to study fine mechanisms and modulation of transmitter release in single active zones.

**Andrei Zefirov, Thomas Benish
Nail Fatkullin, Sergei Cheranov
Roustem Khazipov***

*Department of Physiology,
Kazan Medical University,
Kazan 420012, Russia*

Volcanic risk for Rwandan refugees

SIR — In December 1994, the head of the international humanitarian association "Médecins du Monde", Dr Bernard Granjon, returned from the huge refugee camp just outside the Rwanda-Zaire border, close to the town of Goma, on the southern foot of the large, active volcano Niragongo. Granjon was concerned about possible eruptive activity and sought my advice as a volcanologist who has visited and studied this particular volcano.

In my view, whatever the risk of eruption, there should be no evacuation of such a huge number of people as are living in this camp. I believe this because of long experience as a volcanological forecaster (in charge of major natural hazards for the French government (1981-86) and chair-

man of the national committee for eruptive hazards assessment since 1973).

My advice to ignore the risk from Niragongo is based on two factors: the number of people who could be victims of the potential lava flows; and the number of refugees who would die or contract serious diseases if one million people were to be evacuated from the camp, especially during the rainy season and in the prevailing cold temperatures due to the comparatively high altitudes. While in the camp, the refugees have tents, drinking water and latrines (the lack of these facilities in 1992 and 1993 resulted in cholera epidemics and countless deaths). Even if there were to be swift lava flows (which is by no means certain), transfer of refugees would result in a far greater number of victims than would result from any eruptive event.

As regards the strictly volcanological analysis of the present accumulation of molten lava in the Niragongo crater, the following factors should be considered.

(1) The observed accumulation of lava in the large sink-hole of the volcano will either stop for a long while, as it did from 1982 to 1994, or continue until it flows down the outer slopes of the volcano, either through one of several new fissures that its hydrostatic pressure would open in the sink-hole upper walls or by overflowing the lowest spots of the crater rim.

(2) The observed lava accumulation is 'lava lake' activity, characterized by a convective system of ascending, light, gas-rich and hot magma currents, followed by downwelling of degassed, somewhat cooled and thus denser lava. This phenomenon, which has probably occurred several times during recent geological times, has characterized Niragongo from 1928 to 1977.

(3) Whatever the present situation ('pond' or 'lake' activity), is the whole volcanic pile being pushed upwards by a rise of the whole magma body that feeds the twin volcanoes Niragongo and Nyam-lagira, or does it result from the ascent of Niragongo's magma only?

(4) The observable features of pre-historic Niragongo activity show that a very violent outburst of high-pressure gases could occur, either hurling large volumes of pyroclastics upwards, or then fall with a parabolic trajectory, or propelling them horizontally (or obliquely) by a directed blast. Although such a catastrophe has occurred in the past, the likelihood of a recurrence is low, particularly because the present magma is an exceptionally fluid alkali basanite.

Whatever the risk of any one of these possible hazards, the number of victims of volcanic activity would be at least two to three orders of magnitude smaller than the number that would be claimed by a large-scale displacement of one million refugees.

Haroun Tazieff

*15 Quai de Bourbon,
75004 Paris, France*

1. Rash, J. E., Walrond, J. P. & Morita, M. *J. electron Microsc. Tech.* **10**, 153-185 (1988).
2. Zefirov, A. L., Benish, T. & Fatkullin, N. *Neirofiziolgiya* **22**, 309-318 (1990).
3. Ceccarelli, B., Grohovaz, F. & Hurlbut, W. P. *J. Cell Biol.* **81**, 163-192 (1979).
4. Dreyer, F., Peper, K. & Akert, K. *Brain Res.* **62**, 373-380 (1973).
5. Del Castillo, J. & Katz, B. *J. Physiol., Lond.* **132**, 630-649 (1956).
6. Wernig, A. *Brain Res.* **118**, 63-72 (1976).
7. D'Alonzo, A. Y. & Grinnell, A. D. *J. Physiol., Lond.* **359**, 235-238 (1985).

* Present address: INSERM, Unité 29, Hôpital de Port-Royal 123, Bd de Port-Royal, Paris 75014, France.

† To whom correspondence should be addressed.

On the shoulders of giants

David Hughes

Newton's *Principia* for the Common Reader. By S. Chandrasekhar. Oxford University Press: 1995. Pp. 593. £75, \$130.

ISAAC Newton's *Philosophiæ Naturalis Principia Mathematica* was published in 1687: not only is it the crowning achievement of the seventeenth-century scientific revolution but it is also generally regarded as the most important book in the history of physical science. But no one would call it an easy read. The first edition ran to only about 500 copies. To quote D. T. Whiteside: "In Newton's own lifetime only a handful of talented men, working without distraction at the frontiers of current research, had each in his own way achieved a working knowledge of the *Principia*'s technical content". Even in the 1730s, Voltaire was describing the book as "incomprehensible" and "obscure". George Berkeley actually went so far as to consider the mathematical treatment of motion to be an intellectual abstraction. It is also clear that Subrahmanyan Chandrasekhar, the author of the book under review, agrees with these sentiments. This 1983 physics Nobel prize-winner and emeritus professor of theoretical astrophysics at the University of Chicago does, however, go on to say that he regards the *Principia* as not only unsurpassed but also unsurpassable. To Chandrasekhar, Newton was not "merely a chip off the old block, but the block itself".

Chandrasekhar has set himself the task of overcoming the syntactical problems associated with the often convoluted style that Newton had perforce to adopt in presenting delicate mathematical arguments in continuous prose. In *Newton's Principia for the Common Reader*, he makes considerable efforts to circumvent Newton's propensity for a secretive style. Much in the *Principia* indicates that Newton had a cerebral approach to solving problems. He did not resort to pencil and paper; he held problems in his mind for hours and days and weeks until they had surrendered their secrets. Only then did he dress up the solution for the purpose of exposition.

Newton found that writing down his work was a chore so he often resorted to short cuts and the introduction of phrases such as "it is manifest that", "hence it comes to pass" and "by like reasoning".

Chandrasekhar transforms the Newtonian mathematics into modern idiom and thus makes it much more accessible to what he quaintly refers to as the "common reader". This is not to imply a universal readership: people will need at least the equivalent of a mathematics

degree to understand the proofs. The word 'common' comes from a quote from Dr Samuel Johnson, and in essence it refers to someone with common sense (and a great deal of it to my mind).

Chandrasekhar arranges the proofs in a linear sequence of equations and arguments. The beauty, clarity and economy of Newton's achievements shine through. Chandrasekhar's "personal reflections" are even more enlightening. He spices his book with a running sequence of introductions and commentaries that reveal both an incisive understanding of the way in which Newton's mind worked and also the pleasure that comes from studying and interpreting Newton's greatest book.

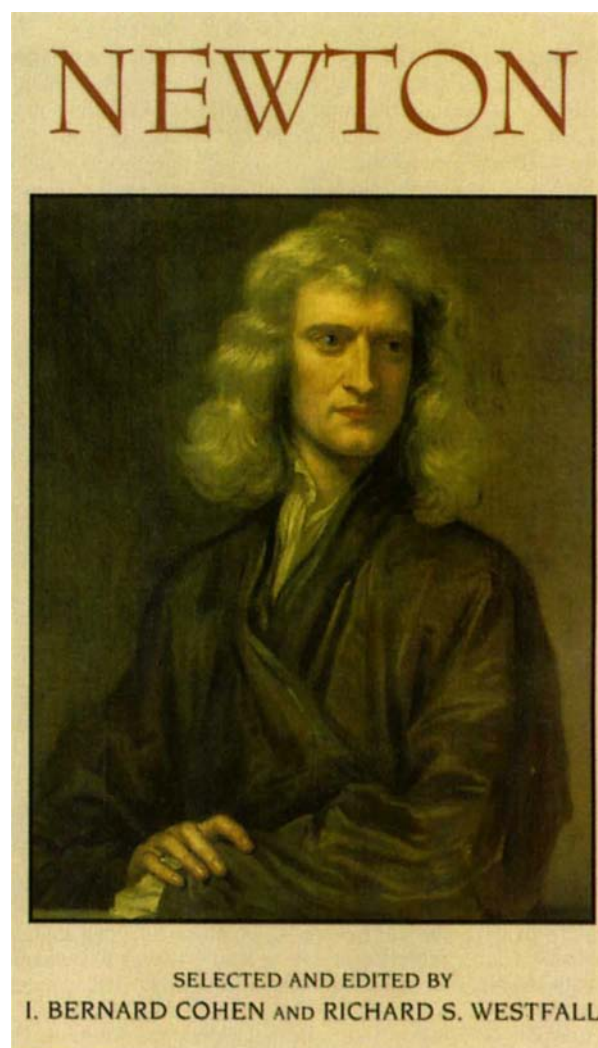
Chandrasekhar does not tackle the

whole of the *Principia* but concentrates on the 150 analytical (that is, purely mathematical) propositions that form the foundation of Newton's deductive presentation of the universal law of gravitation. He stresses that a further attraction of Newton's work was the realization that the dynamical laws were universal too, and that his hypotheses underpinned the concept of absolute space, time and motion. Newtonian gravitation led to the explanation of such disparate phenomena as the tides in the seas, the orbits of comets, the precession of the spin axis of the Earth, the motion of the Moon, the effect of air drag on the descent of bodies and the spin-induced shape of our planet.

The great joy of Chandrasekhar's book is that it repays all the attention one gives it. As one proceeds, everything becomes so much easier to understand. The veil of Newtonian obscurity is lifted and one begins to grasp the extent of Newton's achievement. □

David Hughes is in the Department of Physics, University of Sheffield, Hicks Building, Sheffield S3 7RH, UK.

Newton's thought and achievement are also explored in *Newton* edited by I. Bernard Cohen and Richard S. Westfall (Norton, £6.99 (pbk)). Intended for scholars and students alike, as well as general readers, the book presents a collection of the writings of the great man, together with commentaries spanning more than four centuries by the likes of George Berkeley, Albert Einstein, Colin Maclaurin, D. T. Whiteside, Alfred North Whitehead, John Maynard Keynes and Alexandre Koyré. The cover, shown here, reproduces Sir Godfrey Kneller's portrait of Newton, painted in 1689, when Newton was 46 years old.



Bonds of friendship

Brian Pippard

Fritz London: A Scientific Biography. By Kostas Gavroglu. Cambridge University Press: 1995. Pp. 323. £50, \$69.95.

FRITZ London, the son of a professor of mathematics at Bonn, was born into a Jewish family in 1900. He had reached the right age when the discovery of quantum mechanics initiated the great surge of theoretical physics in the late 1920s, and his background was right for the education that prepared him to play a part. London was baptized in his youth, but this was no help when Hitler came to power; in 1933, he left Germany for good. After some years in Oxford and Paris, he settled at Duke University, North Carolina, where he spent 15 years until his early death in 1954.

London's wanderings play a central role in Kostas Gavroglu's account. With no great facility for languages, and a leaning to the patient rigour of former German scholarship, he did not settle down easily and made few close friends. Lindemann, also German-born, who had built a strong team of refugee physicists at the Clarendon Laboratory, could give only limited support, and he felt himself excluded from the closed shop of collegiate Oxford. Others in the team — Simon, Kurti, Mendelssohn among them — fitted in better and were able to settle down there.

The Paris years were happier, enlivened by the friendship of Langevin, Perrin and Joliot, whose sympathy with communism was no handicap to a liberal intellectual and victim of the Nazis. But in 1939, anxious at the prospect of war, London accepted an offer from Duke University, not then the most congenial place for one of his temperament; the racialism he found there was repugnant. The Austrian physicist Paul Ziesel was only too well aware of the problem from his own parents' tragedies, and expressed his understanding with the remark: "He wanted to be accepted, yet he could not hide his contempt". In the McCarthy era, both were interrogated about their attitudes to communism, but while London was given clearance as a possible consultant at Los Alamos, Ziesel was forced to accept paid leave from his post at Connecticut University and excluded from the campus. In the event, London's death put an end to the projected consultancy. He had been ill for some years and unfit for more than the mildest exertion, and his original intellectual vigour had faded. The two volumes of *Superfluids* (1950 and 1954) absorbed most of his residual energy, for all that they systematize, with little enlargement, ideas that had been his main occupation ever since 1935.

London's innovations mostly belong to the period 1927–39. The first of real note, in collaboration with Heitler, showed how quantum mechanics accounts for the binding of the hydrogen molecule — an exciting and unexpected

discovery. They had set out to elucidate the van der Waals force between two hydrogen atoms and were astonished to find they had discovered homopolar bonding. The overlap of atomic wave-functions was the key mechanism; at greater distances, where there is no overlap, the polarization of one atom by the fluctuating dipole moment of the other is responsible for van der Waals attraction. This was sorted out by London a few years later. Meanwhile, he had attempted to extend the theory of homopolar bonds to other molecules, but was overtaken by the more intuitive procedures of such as Mulliken and Pauling.

Characteristically, London did not favour such methods. He preferred precise formulation, even in situations where he recognized that their complexity demanded abandoning a severe reductionism in favour of what came to be called phenomenological theories. This paid a handsome reward when in 1935 he devised the equations for the supercurrent that have since been central to studies of superconductivity. The original inspiration may have been Fritz London's, but many of the consequences were recognized by his brother Heinz, seven years his junior. It was some years before anyone else saw so clearly the significance of these ideas, and the brothers rightly share the credit.

Fritz was less successful with the apparently related problem of superfluidity in liquid helium, whose origin he attributed to the Bose–Einstein condensation. When Tisza used the notion to construct his two-fluid model, London became a strong supporter against Landau's alternative theory — undoubtedly urged in this direction by Landau's seemingly cavalier disregard of rigour. In the end, measurements of second-sound velocity gave victory to Landau. It is easy, some 50 years later, to appreciate that Landau's physical instinct was of a higher order than London's, but the problems were new and puzzling, and few could then see where the truth lay.

Even if, as I now judge, London was not quite in the top class of physicists, his ideas were stimulating and some have undoubtedly found a permanent place in the textbooks. Gavroglu has given proper attention to the physics, and has done particularly well to use it as a framework for the story of a troubled life. I am personally grateful for the full-length picture of a man from whom I learnt many things, who always had time for serious discussion and whose welcoming smile is among my happiest of memories. □

Sir Brian Pippard, emeritus professor of physics, is at the Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, UK.



PALMS — the quintessential tropical plants — are of immense ecological and economic value. The South American species shown here have a myriad of uses: the fruits provide a popular beverage, which in turn can be boiled down to produce a clear oil; trunks are used in construction; leaves are woven into baskets; and leaf sheath fibres are used for blowpipe darts. The photographs appear in the comprehensive and detailed *Field Guide to the Palms of the Americas* by A. Henderson, G. Galeano and R. Bernal. Princeton University Press, \$75, £50.

In search of Neanderthals

Robin Dennell

Honor Among Thieves: A Zooarchaeological Study of Neandertal Ecology. By Mary C. Stiner. *Princeton University Press*: 1995. Pp. 447. \$69.50, £46.50.

Mousterian Lithic Technology: An Ecological Perspective. By Steven L. Kuhn. *Princeton University Press*: 1995. Pp. 209. \$49.50, £39.50.

JUST occasionally, books on palaeo-anthropology can be truly exciting. Here are two examples that break new ground in enhancing our understanding of Neanderthals and have important implications for how we study the behaviour of past populations. Both deal with the ancient coastal region of Latium, Italy, during the early Upper Pleistocene (35,000–100,000 years ago (35–100 kyr)). The area is known mainly for its curious 'Pontian Mousterian', made largely of beach pebbles, and for Monte Circeo, an alleged (and undoubtedly spurious) site of Neanderthal ritual.

One of the biggest questions in palaeoanthropology concerns the emergence of modern humans. In Europe, the problem is tied to two events that happened 30–40 kyr. First, Neanderthals either became extinct or evolved into fully modern humans (either way, there is no sign of them after 30 kyr). Second, and perhaps coincidentally, there is a transition from the Middle Palaeolithic period to the Upper Palaeolithic period around 35 kyr, marked primarily by the predominance of blade stone tools over flake tools, and the appearance of bone projectile points, hearths and art. The transition has (regrettably) dictated the course of most discussion on the emergence of modern humans. It is regrettable because it focuses attention only on the end of Mousterian culture (the final part of the Middle Palaeolithic) and not on its previous 100,000 years; there is no reason why this transition should have resulted in changes in anatomy, genetic make-up, subsistence, social behaviour or indeed anything except a limited range of technological changes; it also leads to the Mousterian being seen as monolithic rather than as a phenomenon with its own internal history and dynamics. This is grossly unfair to the earlier Neanderthals, who deserve to be studied in their own right as an extinct form of humanity.

The Mousterian itself is complex and

has been studied in a bewildering number of ways. Continental archaeologists have tended to concentrate on stone tool technology and typology and, more recently, on the "chaîne opératoire" (the sequence in which flakes and blades were removed from the core of stone) and the life histories of artefacts and their function. British and US workers have relied more on geographical, geological and anthropological perspectives. Studies have ranged from individual sites to local regions (such as the Périgord in south-west France) and even larger climatic areas. Researchers studying animal remains from Upper Pleistocene sites (usually caves) have looked for changes or patterns in hunting strategies, using various methods that often predetermine the conclusions. Disagreements are rife, arising from differences in nationalities, professions, training and

prime-age animals — one of our unique adaptations as a predator. Her analysis is provocative in that she suggests that this change took place long before the Middle–Upper Palaeolithic transition 35 kyr and was accomplished without the use of bone- and stone-tipped projectiles.

Kuhn's concern is with the way in which Neanderthals maintained a supply of raw materials and artefacts as part of their survival strategy. His main findings match Stiner's. Before 55 kyr, Neanderthals were provisioning individuals with large renewable flake tools that could be re-sharpened while on scavenging trips; thereafter, the emphasis was on provisioning places in preparation for hunting prime-age animals, with tool-makers switching to the more economical technique of working cores and producing fresh blanks from local pebbles rather than re-sharpening old tools.

Before and after 55 kyr, Neanderthals obtained meat from different body parts and used different methods and toolkits for doing so. Kuhn shows that the changes were not accompanied by the appearance of new hunting items or methods of making stone tools; these changes would thus escape many continental researchers fixated on technologies or typologies.

The implications of all this are fascinating but unclear. Locally, the primary driving force behind the changes seems to be climatic: after 55 kyr, cave sites became further away from the coast owing to lowered sea levels. Because these climatic changes were global, other Neanderthal groups may have responded similarly. The books

also raise important questions. Did Neanderthals use a wider range of techniques for obtaining meat from large animals than humans after 35 kyr? How and why did they ambush large animals without using hafted projectiles? Why didn't they often use fire for processing meat, as was common after 35 kyr? And, of course, why did they die out after being successful for so long?

These two remarkable books are landmark contributions that are outstanding in their theoretical and methodological rigour. Although they will remain largely inaccessible to nonspecialists, they should become role models for future related work. As the authors show so clearly, Neanderthals (and their predecessors) are fascinating in their own right, and not merely because the last of them may (or may not) have evolved into us. □

Robin Dennell is in the Department of Archaeology and Prehistory, University of Sheffield, Sheffield S1 4ET, UK.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Early humanity — artist's reconstruction of Neanderthal life.

inclinations of the researchers.

Both authors are well versed in evolutionary ecology and anthropology, aware of the need for testable hypotheses and prepared to treat the Mousterian as a phenomenon in its own right and not simply as a possible precursor of the Upper Palaeolithic. They examine the same sites (Mary Stiner concentrates on faunal remains, Steven Kuhn on artefacts); they bravely base their studies on excavations of several decades ago, when recording and collecting techniques were usually less stringent than today's; and they fully incorporate the past 20 years of research into their methodology.

Stiner looks at evolutionary ecology in terms of guilds of human and nonhuman predators (notably hyena and wolf) competing for the same prey (ungulates). Two patterns of Neanderthal subsistence emerge: the first, more than 55 kyr, is marked by the scavenging of head parts, perhaps to obtain fat in lean seasons; the second, later on, shows the ambushing of

Immortal longings

Leonard Hayflick

Ageing: A Biomedical Perspective. By Denis Bellamy. Wiley: 1995. Pp. 410. £35.

Understanding Ageing. By Robin Holliday. Cambridge University Press: 1995. Pp. 207. £35, \$64.95 (hbk); £14.95, \$24.95 (pbk).

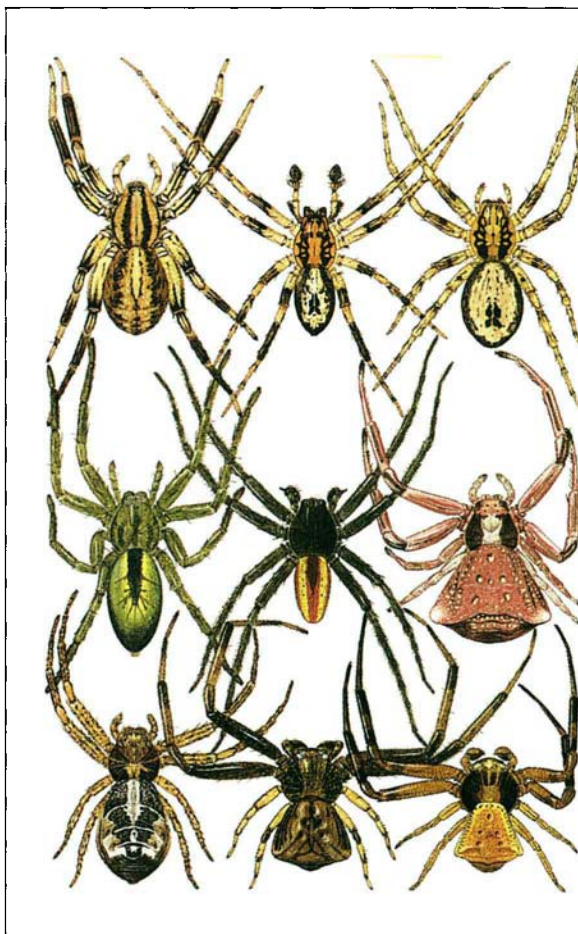
AGEING is an aberration of civilization. Its extreme manifestations affect only humans or those animals we choose to protect. We have learned how to eliminate most causes of youthful death, allowing most people in developed countries to live long enough to experience ageing — a phenomenon that, teleologically speaking, we were never intended see.

Remains of old prehistoric humans have rarely been found. Similarly, feral animals do not usually live long enough to age: after reproductive success, they incur physiological losses that increase their vulnerability to predation, disease or accidents. When feral animals do reach old age, it is often because humans have disturbed their ecological niche.

Because we have learned how to prevent most early deaths, ageing is happening on a scale never before experienced by humans. Not only are the numbers of old people increasing in developed countries but life expectation at almost every age is advancing. The political fallout from these events, together with the resulting availability of significant research funds for the first time, is one of the reasons why biogerontology has emerged suddenly from a scientific backwater to become a fashionable pursuit.

Denis Bellamy makes a heroic and largely successful effort to present a balanced survey of most of the salient beliefs that underpin the field. This is not an easy task. The complexity of biogerontology is staggering because it embraces events from conception to death, from the molecular level to the whole animal, and it even encompasses entire populations.

Robin Holliday offers a superb overview with emphasis on the cellular and molecular biology of ageing — an aspect today enjoying immense popularity. Excitement has been sparked by the twin discoveries of cell-cycle inhibitors and the shortening of the telomeric ends of chromosomes as cells age both *in vitro* and *in vivo*, marking what may be the first demonstration of a molecular clock. Holliday advances several bold personal views, including his belief that the biological basis of ageing in its broad outline is now well understood; it is only the fine detail that remains to be revealed. Despite his compelling arguments, others may not be as optimistic, including Bellamy who prefers



AN array of spiders from the *Collins Field Guide to Spiders of Britain and Northern Europe* by Michael J. Roberts. This compact book covers 450 species, all thoroughly illustrated, and is packed full of information on the structure and biology of spiders, as well as their courtship, hunting and web-making behaviour. The author also describes a clever new method for restraining live spiders that is simplicity itself. HarperCollins, £14.99.

to examine the field critically and in encyclopaedic dimensions.

Holliday claims that the anatomy and histology of many tissues and organs are incompatible with indefinite survival. Yet some biogerontologists have argued that animals that do not reach a fixed size at maturity do not show obvious age changes.

The Galapagos tortoise, sharks and many other fish are examples, as is the American lobster: a 55-year-old specimen weighing 44 pounds has been observed to close its claw as fast as a juvenile. In animals that, like humans, reach a fixed size at maturity, reaction time increases with age. Animals that do not age are not immortal. They eventually succumb to a steady risk of death from predation, disease or accidents.

Both authors discuss ageing as a phenomenon distinct from, and similar to, a disease process. The resolution of this dilemma is important. If ageing is a disease, then we can assume that its remedy is both possible and desirable. If it is not, then perturbing the process may be impossible or, more importantly, undesirable. Unlike ageing, no disease occurs universally and only after reproductive maturity, nor does any disease cross virtually every species barrier. Finally, most age changes are not usually thought of as pathologies. No one has ever died of grey

hair or wrinkled skin.

Both authors present cogent arguments that address the fundamental question in this field: why does ageing occur? A compelling response is that ageing, as distinguished from longevity determination, is a by-product of natural selection, the forces of which diminish greatly after reproductive maturity. Species survival is best guaranteed by selecting individuals with the greatest chance of reaching reproductive maturity rather than of achieving longevity.

Survival to reproductive age is increased by selection that favours greater physiological capacity than would be minimally required. How long individuals will then survive depends on the rate of entropy loss, a process characterized by increased vulnerability to disease, predation and accidents. Thus the amount of excess physiological capacity determined by the genome and reached at maturity is species-specific and determines longevity, but only indirectly. There were never enough old feral animals present for natural selection to favour a genetic programme that directly orchestrates longevity or age changes. □

Leonard Hayflick is in the Department of Anatomy, University of California at San Francisco, PO Box 89, San Francisco, California 95497, USA.

Conflict over the age of the Universe

M. Bolte & C. J. Hogan

The ages of the oldest stars in our Galaxy can be estimated by comparing stellar populations in globular clusters to calibrated stellar models. The best data give cluster ages of about 15.8 ± 2.1 Gyr, which conflicts with the age of the Universe estimated from measurements of the Hubble constant and the 'standard' cosmological model of a flat, matter-dominated universe (8–13 Gyr).

THE cosmic abundances of the light elements and the precise black-body character of the microwave background radiation provide strong evidence that the expansion of the Universe began a finite time (t_0) ago, with the Big Bang¹. Determining just how long ago the Big Bang took place is a fundamental question of cosmology, but a difficult one to answer. The current rate of expansion of the Universe is quantified by the Hubble constant (H_0), the ratio of the speed with which distant galaxies appear to be receding from us to their distance. This rate is slowed with time by the mutual gravitational attraction of all the matter in the Universe², and possibly accelerated by the energy density of the physical vacuum, the effect of which is usually quantified as the 'cosmological constant'^{3,4}. By determining precisely H_0 and t_0 , we can investigate the mass density in the Universe (how much dark matter there is), attempt to establish whether the cosmological constant has a finite value, and test the idea that the Universe went through a period of rapid expansion ('inflation') in its early stages.

The standard cosmological model (on grounds of elegance as well as inflation) is the Einstein–de Sitter model, in which the cosmological constant is set to zero and the mass density is such that the Universe is 'flat', expanding at precisely the velocity needed to sustain that expansion indefinitely. In this model, the age of the Universe is simply related to H_0 by $t_0 = \frac{2}{3}H_0^{-1}$. The global value of H_0 , and hence the Einstein–de Sitter age, has long been uncertain; by current estimates^{5–7}, the Universe may be as old as 13 Gyr (for $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$) or as young as 8 Gyr (for $H_0 = 80 \text{ km s}^{-1} \text{ Mpc}^{-1}$).

An independent constraint on t_0 comes from studies of stellar evolution, since the Universe must be older than its oldest stars. But herein lies the problem. The ages of the oldest stars in our Galaxy are estimated to be about 16 Gyr, significantly older than the Einstein–de Sitter age for current estimates of H_0 . If the Einstein–de Sitter model is to remain the cosmological model of choice, we must then re-examine the possible errors in our model-derived estimates of the ages of the oldest stars.

Here we look at the procedure by which stellar ages are determined to see how much scope there is for reducing these estimated ages. We find that a significant revision to the ages of the oldest stars is improbable and conclude that this age test rules out a flat, matter-dominated Universe.

Estimating stellar ages

The oldest known stars lie in globular clusters (Fig. 1). Reflecting this antiquity, their abundance of elements heavier than helium (the 'metallicity') is extremely low. The Big Bang produced hydrogen, helium and trace amounts of lithium, but all heavier atoms have been created as a result of thermonuclear burning in the centres of stars. These heavy elements are distributed throughout space by the winds from evolved stars and, particularly in the early Universe, as a result of supernovae; therefore on average, the earlier a star formed, the less heavy elements it contains, because the gas from which it formed had not been significantly enriched with these heavy elements.

Stars provide a relatively accurate chronometer because their structure, a delicate quasi-equilibrium where the inward force of

gravity is balanced by the pressure of the gas heated by compression and kept heated by nuclear fusion in the core, changes as the core uses up the hydrogen available for burning. After spending most of its life at an essentially constant surface temperature and luminosity (on the 'main sequence'), a star evolves rapidly to a new equilibrium phase with a lower surface temperature and higher luminosity (the giant phase, of which there are many types). Plotting the 'colours' of cluster stars against their brightness (a colour-magnitude diagram; Fig. 2), a distinctive bend in the distribution clearly marks the transition from the main-sequence to giant-branch phases of evolution. This bend in the colour-magnitude plane is called the main-sequence turnoff (MSTO). After correction for interstellar 'reddening' (the property of grains to scatter blue light more effectively than red light), the colour of a star is used to derive its effective surface temperature (T_{eff}) and, once the distance to a cluster is determined, the brightness of a star gives its luminosity (M_{bol}). The key to using stars as clocks is that for fixed composition, the main-sequence surface temperature, luminosity and lifetime of a star are all functions of its mass alone. By deriving the surface temperature and luminosity of the cluster stars at the MSTO we can use stellar models to derive their mass and main-sequence lifetime. Since these stars are just completing their main-sequence phase of evolution, we have derived the age of the cluster.

Although the details of this procedure are complicated, the individual steps are now well calibrated and realistic estimates of the errors have been made. This does not mean that the method is foolproof, but any modification needs to pass several calibration tests.

Stellar models and calibrations

We live at a convenient time in the evolution of the Universe: stars now near the MSTO of globular clusters (approximately one solar mass, $1 M_{\odot}$ and 'metal-poor') are those for which our models are the most reliable^{8–12}. Stars of around $1 M_{\odot}$ are intrinsically less complicated than more or less massive stars. They have radiative cores, relatively simple atmospheres and, throughout most of their interiors, the equation of state does not deviate far from that of a perfect gas. Stars more massive than about $1.5 M_{\odot}$ have core convection and 'overshooting' (the tendency of convective momentum to propel gas beyond its equilibrium depth)—phenomena that have yet to be completely understood. For stars less massive than about $0.5 M_{\odot}$ the equation of state becomes increasingly complex, and the transformations from the $T_{\text{eff}}-M_{\text{bol}}$ plane of stellar interior theory to the equivalent observational quantities of colour index and absolute magnitude in a particular bandpass are uncertain owing to the difficulties of accurately calculating model atmospheres for relatively low temperatures where molecules are plentiful.

Model predictions are insensitive to current uncertainties in stellar 'physics' (interaction cross-sections for fusion reactions, and the equation of state at different radii), and to the details of the numerical algorithms used to solve the stellar structure equations^{8,11}. Recent improvements include the use of non-solar abundance ratios, most importantly the over-abundance of oxy-



FIG. 1 Globular clusters are discrete clusters of a million or more stars, which are held together by the mutual gravitational attraction of all the stars. Approximately 200 globular clusters orbit the centre of the Milky Way. This cluster, M92, contains stars which are among those having the lowest abundance of elements heavier than helium of any stars observed; the abundance of iron in these stars is a factor of 150 less than in the Sun. Precise photometry of M92 provides our best estimate of the oldest stellar ages.

gen compared to scaled solar values for metal-poor systems (which has reduced the inferred ages of clusters by about 15%)^{13,14}, hydrostatic effects and mixing due to internal rotation, and the effects of helium diffusing downwards through low-mass stars—another ~5% reduction^{15–17}. Remaining shortcomings of the models (the parametrized treatment of convection by the so-called ‘mixing-length’ theory, and an inadequacy in calculating opacities in intermediate temperature regimes of about 10^5 K) primarily affect the calculated radii and therefore effective temperatures and colours of model stars. There is recent progress in both areas^{18,19} but model luminosities and the mass–lifetime relations, particularly for the most metal-deficient models, are insensitive to the treatment of convection and opacity.

The main reason to believe that systematics are well controlled is that models pass several calibration tests. With very accurately known values for its age, luminosity, effective temperature and abundance parameters (although helium is not measured

directly), the Sun provides a precise test of stellar models for stars near $1 M_{\odot}$. At relatively high metallicity, solar models suffer much more from the inadequacy of current opacity tables than do models appropriate to globular-cluster stars. Although the solar-neutrino problem has been the source of significant consternation to stellar modellers, it seems likely that the resolution of that mystery will not be found in changes to the solar standard model, but rather in our understanding of neutrino physics²⁰. (Note that neutrino losses in any case only contribute to a few per cent of the solar luminosity). Certain eclipsing binary stars (whose orbits are thereby known to be viewed nearly edge-on), for which velocity curves can be derived for both stars in the binary system, allow very accurate determinations of both the individual stellar masses and the distance to the system. Studies of the known handful of such objects have provided strong confirmations of the mass–luminosity and colour–luminosity relations predicted by the models, even off the main

FIG. 2 a, The $[\text{Fe}/\text{H}] = -2.26$ zero-age main sequence (squares) formed by the solar-neighbourhood subdwarfs. These are the subdwarfs with trigonometric parallax measures for which the errors are 30% or less and $[\text{Fe}/\text{H}] \leq -1.3$. Small corrections determined by interpolation in the models of Bergbusch and Vandenberg¹⁴ have been made to the colours of the subdwarfs with $[\text{Fe}/\text{H}]$ values other than -2.26 , to bring them to a ‘mono-metallicity’ sequence. The solid lines are models for $[\text{Fe}/\text{H}] = -2.26$, $Y = 0.235$, $[\text{O}/\text{Fe}] = +0.7$ and ages (in descending order) of 12, 14, 16 and 18 Gyr. b, As a with the addition of the binned data from M92 plotted as filled circles. The M92 sequences have been corrected for reddening (0.02 mag in $E(B-V)$) and adjusted vertically to give the best fit to the subdwarf sequence between the colours 0.55 and 1.1.

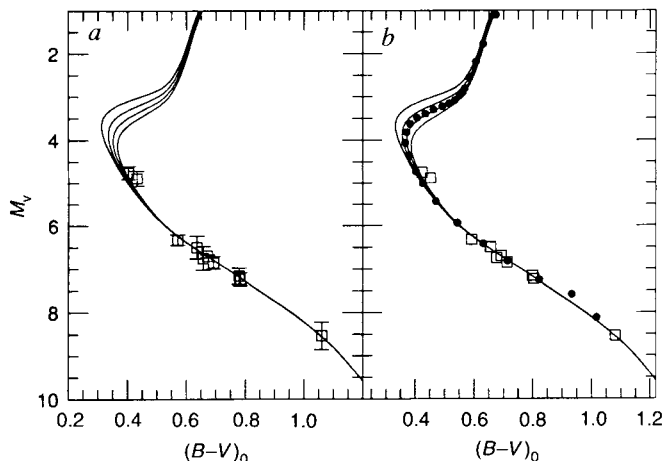


TABLE 1 Typical error budget for age determinations at $[\text{Fe}/\text{H}] = -1.6$

Input parameter	Typical errors	$(\delta \text{ age}/\text{age}) \times 100$	
		Colour (MSTO)	M_V (MSTO)
Y	0.02	5%	2%
$[\text{Fe}/\text{H}]$	0.2	16%	6%
$[\text{O}/\text{Fe}]$	0.2	13%	6%
α	0.3	8%	<1%
D	25%	—	22%
$E(B-V)$	0.02 mag	16%	2(8)%*
$\delta V, \delta(B-V)$	0.02 mag	16%	2(8)%*

Column 1 shows the input parameter, column 2 the typical errors in each parameter for nearby clusters, column 3 the fractional error in the age estimate as a result of the parameter uncertainty if ages are determined by comparing models to main-sequence turnoff (MSTO) colours and column 4 the fractional error in derived ages if MSTO luminosities are used to derive ages⁹. Symbols used: Y , helium mass fraction; $[\text{Fe}/\text{H}]$, decimal logarithm of iron abundance relative to solar; $[\text{O}/\text{Fe}]$, oxygen abundance relative to solar; α , 'mixing length' parameter for convection; D , distance in linear units; $E(B-V)$, colour excess due to intervening dust; $\delta V, \delta(B-V)$, typical uncertainties in photometric zero-points. The asterisk indicates quantities that depend on the method used to derive cluster distance.

sequence^{21,22}. Most convincing is the case of AI Phoenicis, a double-lined eclipsing system where one star is a subgiant and the other just evolved above the zero-age main sequence. The models successfully predict the same age for the two components with a precision²³ of $\sim 1\%$ —a remarkable coincidence if ages are systematically wrong, as they would have to be wrong by exactly the same amount for two stars of very different mass.

Finally, the models predict accurately the colour–luminosity relation (including the metal dependence^{10,14}) for the unevolved metal-poor low-mass stars (which have just joined the main sequence) in globular clusters and, even better, in the local versions of globular-cluster main-sequence stars, the so-called subdwarfs. These are members of the Galactic halo which happen to be near enough to the Sun to have measurable trigonometric parallaxes. The nearby subdwarfs are also thereby the linchpin of the globular-cluster distance scale, and yield a distance calibration confirmed by RR Lyrae stars in the Large Magellanic Cloud—the foundation of the extragalactic distance scale^{24,25}.

Age estimates and errors

To estimate an age we need to determine the mass of those stars at the MSTO given the observed apparent brightness and colour of the MSTO. The observed magnitude and colour of the MSTO is converted to absolute magnitude M_V and colour $(B-V)_0$ with knowledge of the distance to, and reddening towards, the cluster. These quantities are converted to the magnitude of the total energy produced in fusion reactions in the core M_{bol} and the effective temperature of the stellar photosphere T_{eff} through the use of a model stellar atmosphere for the appropriate abundances ($[\text{Fe}/\text{H}]$, $[\text{O}/\text{Fe}]$ and helium mass fraction Y) and convolution of the calculated spectrum with the response function of the filter system, telescope and detector. The values of M_{bol} and T_{eff} for stars at the MSTO then give the mass and lifespan when used in a stellar-interiors model with the correct abundances. Although this is a fairly complex procedure with a formidable list of input parameters (as discussed above), for the case of metal-poor, low-mass stars, the stellar-interiors and stellar-atmosphere models are well understood. The important questions are: how do uncertainties in determining the input quantities propagate through the determination of cluster age, what is the best age estimate and how large are the random and systematic errors in this age?

We can use the models to derive a table of the influence of the various input parameters on age estimates. Table 1 summar-

izes the results for the cases of estimating ages by MSTO colour and MSTO luminosity.

It is clear from Table 1 that, in addition to stellar colours being the most uncertainly predicted quantity in the stellar models, uncertainties in virtually every input parameter contribute significantly to the uncertainty in derived age when using the colour of the MSTO as an age indicator, which is why turnoff colours are to be avoided as an age diagnostic. On the other hand, if the MSTO luminosity is used to derive cluster ages, the only significant source of error is in cluster distance determinations. This is the crux of the age issue.

The local subdwarfs can directly tie cluster distances (and ages) to the most reliable distance scale in astronomy, that provided by trigonometric parallax, or triangulation using the Earth's orbit as a baseline^{26–28}. For each subdwarf we therefore have a 'standard candle' at a particular metal abundance and colour which, when compared to main-sequence stars of the same colour in any cluster, gives the distance to the cluster from the inverse square law of radiation. Specifically, after correcting for interstellar reddening the colours and magnitudes of the stars on the cluster unevolved main sequence, the vertical shift required to match the cluster main sequence to the sequence defined by the local subdwarfs gives the distance to the cluster. (For comparison with any one cluster, each subdwarf's observed colour must also be 'corrected' to give the colour that star would have at the metallicity of the cluster.)

The sample used for the past 10 years comprises only eight metal-poor stars with distances and M_V values known to 30% or better. Moreover, until recently the nearby subdwarfs, even corrected to a mono-metallicity sequence, did not form a particularly tight fiducial for deriving distances^{29–31}, contributing most of the uncertainty ($\sim 15\%$) in the distance modulus given in Table 1. But with the availability of the newest compilation of trigonometric parallax measures in 1991³² and with new measurements of $[\text{Fe}/\text{H}]$, V and $B-V$ for some of the local subdwarfs, it appears that much of the scatter seen in the previous subdwarf sequence was due to observational inaccuracies. Figure 2a shows the sequence of subdwarfs with distance errors $<35\%$ and colours corrected to $[\text{Fe}/\text{H}] = -2.26$ (squares); superposed are two model isochrones from Bergbusch and Vandenberg¹⁴ for this metallicity and ages of 12, 14, 16 and 18 Gyr. The model colours were uniformly adjusted by 0.015 mag (to the red) to match the colour of the subdwarf with the largest parallax, HD 103095 (Groombridge 1830)³², but other than that there has been no further adjustment of model parameters to fit the subdwarf data. The agreement not only encourages confidence in the models, but provides a precise fix on cluster distances.

The age of M92

The most accurate photometry for a globular cluster with high galactic latitude (to avoid reddening uncertainties) and low metallicity (where the models are best and the clusters possibly oldest) remains the study of M92 by Stetson and Harris³¹. Figure 2b shows as filled circles the mean position of the locus of points for the M92 stars after correcting for reddening (0.02 mag in $E(B-V)$) and making a vertical registration based on a least-squares fit of the unevolved M92 main sequence to the subdwarfs of Fig. 2a. This gives a distance of 8.5 kpc to M92 and puts the bluest point of the MSTO region at $M_V = 4.0$. Superposed on the M92 data are isochrones from ref. 14 for $[\text{Fe}/\text{H}] = -2.26$, $[\text{O}/\text{Fe}] = +0.7$, $Y = 0.235$ and for 12, 14, 16 and 18 Gyr. The agreement between the isochrone shapes and the M92 sequences is impressive, but the most reliable age estimate comes from comparing the observed turnoff, M_V of the MSTO for M92, with the model-derived M_V –main-sequence lifetime relation for the metallicity of M92. Consistent with the 'fit' in Fig. 2b, the age of M92 based on M_V (MSTO) = 3.95 from the $[\text{Fe}/\text{H}] = -2.26$ models is 15.8 Gyr. The random errors as in Table 1 can be propagated to give a standard deviation $\sigma = 2.1$ Gyr with the uncertainty shared equally between the distance error and the

uncertainty in $[\text{Fe}/\text{H}]$. This value for σ does not include any systematic error in the models.

It is a much easier task to determine relative cluster ages, and it has been shown convincingly that the age difference between M92 and essentially all of the other nearby clusters with $[\text{Fe}/\text{H}] \lesssim -2$ is <0.8 Gyr (ref. 33), and that M92 is probably a good representative of the metal-poor halo of the Galaxy as far as age is concerned.

Although there are no other comparably precise methods of estimating ages for objects as old as globular clusters, the cooling ages of white dwarfs in the disk of the Milky Way provide an independent check on the globular-cluster ages. The disk of the Milky Way probably formed much later than the globular clusters, so the white-dwarf ages determined by models of their cooling (between 7.5 and 11 Gyr; ref. 34) are consistent with the globular-cluster ages. The white-dwarf ages also coincide with the ages for the oldest 'open clusters' in the disk of the Galaxy³⁵.

The age of the Universe

The expansion age of the Universe, defined as the inverse Hubble constant $H_0^{-1} = 9.78 \times 10^9 (H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1})^{-1} \text{ yr}$, sets the timescale of all cosmological models. It is simply the relation between the expansion velocities of distant galaxies and their distances, $v = H_0 r$, and therefore even an approximate agreement with stellar ages (as seen) confirms the basic premise of a Big Bang at the beginning of an evolving Universe, as the expansion age is completely independent of the ages of the stars.

The increasingly reliable measurements of the Hubble constant seem to be converging on a value around 70 to 80 $\text{km s}^{-1} \text{ Mpc}^{-1}$ (refs 5, 6, 36–41). Although some dispute remains about the calibration of type Ia supernovae (which have been used to argue for a smaller value of H_0) as standard candles^{42–46,50}, and about the structure and depth of the Virgo cluster⁵¹, the formal uncertainty of many determinations is $\chi \pm 10 \text{ km sec}^{-1} \text{ Mpc}^{-1}$. We probably know H_0 to about the same precision—something

like $\pm 25\%$ at the 2σ level—that we know the ages of the oldest stars.

The current data already significantly constrain models that had higher expansions rates in the past, particularly the 'standard' Einstein–de Sitter cosmology (with $\Omega = 1$, and zero cosmological constant) which predicts an age of $t_0 = (2/3)H_0^{-1}$; this differs from the stellar ages by about 3σ for $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Even for $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$, the value favoured by some groups using type Ia supernovae as distance indicators^{42–44}, the Einstein–de Sitter cosmology predicts an age of only 13 Gyr, which is still hard to reconcile with the ages of the oldest stars. If we allow 1 Gyr for the first stars to form, they should be no older than 12 Gyr now, which is about 2σ younger than M92. Within the context of general relativistic cosmological models, those with low mass density ($\Omega < 1$) or models dominated by a cosmological constant would now seem to be favoured by the data. Unless there is some important missing ingredient in the models for stars around $1 M_\odot$ —and such an ingredient would have to leave the mass–luminosity and luminosity–colour relations essentially unchanged—we conclude that the age test already rules out the standard cosmological model.

It is improbable that in the near future a strong case could be made on the basis of the age test for a nonzero cosmological constant, or that we will be able to differentiate between flat and open cosmological models, as realistic errors on t_0 and H_0 will still permit some uncertainty. Further progress will require 'non-local' tests of the change with time of the expansion rate of the Universe, which may be possible using the distances to very distant galaxies (perhaps using supernovae as standard candles⁴⁷) or other methods that are sensitive to global geometry^{48,49}. Perhaps these studies will offer clues to resolving the current conflict. $\square$

M. Bolte is at Lick Observatory, University of California, Santa Cruz, California 95064, USA. C. J. Hogan is at the Astronomy and Physics Department, University of Washington, Seattle, Washington 98195, USA.

- Peebles, P. J. E., Schramm, D. N., Turner, E. L. & Kron R. K. *Nature* **352**, 769–776 (1991).
- Coles, P. & Ellis, G. *Nature* **370**, 609–615 (1994).
- Carroll, S. M., Press, W. H. & Turner, E. L. A. *Rev. Astr. Astrophys.* **30**, 499–532 (1992).
- Weinberg, S. *Rev. mod. Phys.* **61**, 1–23 (1989).
- Freedman, W. L. et al. *Nature* **371**, 757–762 (1994).
- Pierce, M. J. et al. *Nature* **371**, 385–389 (1994).
- Hogan, C. J. *Nature* **371**, 374–375 (1994).
- Rood, R. T. in *Astrophysical Ages and Dating Methods* (eds Vangioni-Flan, E. et al.) 313 (Frontieres, Gif sur Yvette, 1990).
- Renzini, A. in *Observational Tests of Inflation* (eds Banday, T. & Shanks, T.) 131 (Kluwer, Dordrecht, 1991).
- Demarque, P., Deliyannis, C. P. & Sarajedini, A. in *Observational Tests of Inflation* (eds Banday, T. & Shanks, T.) 355 (Kluwer, Dordrecht, 1991).
- VandenBerg, D. A. in *The Formation and Evolution of Star Clusters* 183 (ed. Janes, K.) (ASP Conf. Ser. Vol. 13, 1991).
- VandenBerg, D. A. *Astrophys. J.* **391**, 685–709 (1992).
- Salaris, S. C., Chieffi, A. & Straniero, O. *Astrophys. J.* **414**, 580–600 (1993).
- Bergbusch, P. A. & VandenBerg, D. A. *Astrophys. J. Suppl. Ser.* **81**, 163–220 (1992).
- Deliyannis, C. P., Demarque, P. & Kawaler, S. D. *Astrophys. J. Suppl. Ser.* **73**, 21–65 (1990).
- Proffitt, C. R. & VandenBerg, D. A. *Astrophys. J. Suppl. Ser.* **77**, 473–514 (1991).
- Chaboyer, B. & Demarque, P. *Astrophys. J.* **443**, 510 (1994).
- Chieffi, A., Straniero, O. & Salaris, M. *Astrophys. J.* **445**, L39–L42 (1995).
- Canuto, V. M., Minotti, F. O. & Schilling, O. *Astrophys. J.* **425**, 303–325 (1994).
- Bahcall, J. N. & Bethe, H. A. *Phys. Rev. D* **47**, 1298–1301 (1993).
- Andersen, J. *Astr. Astrophys. Rev.* **3**, 91–126 (1991).
- Swenson, F. J., Faulkner, J., Rogers, F. J. & Iglesias, C. A. *Astrophys. J.* **425**, 286 (1994).
- Andersen, J. et al. *Astr. Astrophys.* **196**, 128 (1988).
- Walker, A. R. *Astrophys. J.* **390**, L81–L84 (1992).
- Carney, B. W., Storm, J. & Jones, R. V. *Astrophys. J.* **386**, 663–684 (1992).
- Sandage, A. R. *Astrophys. J.* **162**, 841 (1970).
- Laird, J. B., Carney, B. W. & Latham, D. W. *Astr. J.* **95**, 1843 (1988).
- Ryan, S. G. *Astr. J.* **104**, 1144–1155 (1992).
- Fahlman, G. G., Richer, H. B. & VandenBerg, D. A. *Astrophys. J. Suppl. Ser.* **58**, 225 (1985).
- Bolte, M. *Astrophys. J.* **319**, 760–771 (1987).
- Stetson, P. B. & Harris, W. E. *Astr. J.* **96**, 909 (1988).
- van Altena, W. F., Lee, J. T. & Hoffleit, E. D. *The General Catalogue of Trigonometric Parallaxes: A Preliminary Version* (electronic preprint, Yale University, 1991).
- VandenBerg, D. A., Bolte, M. & Stetson, P. B. *Astr. J.* **100**, 445–468 (1990).
- Wood, M. A. *Astrophys. J.* **386**, 539–561 (1992).
- Janes, K. A. & Phelps, R. L. *Astr. J.* **108**, 1773–1785 (1994).
- Schmidt, B. P. et al. *Astrophys. J.* **432**, 42–48 (1994).
- Pelt, J., Kayser, R., Refsdal, S. & Schramm, T. *Astr. Astrophys.* (submitted).
- Schild, R. E. & Thomson, D. J. *Astr. J.* (in the press).
- Fukugita, M., Hogan, C. J. & Peebles, P. J. E. *Nature* **366**, 309–312 (1993).
- Jacoby, G. et al. *Publs astr. Soc. Pacif.* **104**, 599–662 (1992).
- van den Bergh, S. *Publs astr. Soc. Pacif.* **106**, 1113–1119 (1994).
- Saha, A. et al. *Astrophys. J.* **425**, 14–34 (1994).
- Sandage, A. et al. *Astrophys. J.* **423**, L19–L22 (1994).
- Saha, A. et al. *Astrophys. J.* **438**, 8–26 (1995).
- Phillips, M. *Astrophys. J.* **413**, L105–L108 (1993).
- Riess, A. G., Press, W. H. & Kirshner, R. P. *Astrophys. J.* **438**, L17–L20 (1995).
- Perlmutter, S. et al. *Astrophys. J.* **440**, L41–L44 (1995).
- Rix, H.-W., Maoz, D., Turner, E. L. & Fukugita, M. *Astrophys. J.* **435**, 49–54 (1994).
- Mao, S. & Kochanek, C. S. *Mon. Not. R. astr. Soc.* **268**, 569–580 (1994).
- Hamuy, M. et al. *Astr. J.* **109**, 1–13 (1995).
- Mould, J. R. et al. HST key project preprint (1995).

ACKNOWLEDGEMENTS. This work was supported by the US NSF and NASA.

Runaway telomere elongation caused by telomerase RNA gene mutations

Michael J. McEachern^{*} & Elizabeth H. Blackburn^{*†}

^{*} Department of Microbiology and Immunology, Box 0414, [†] Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143, USA

The ribonucleoprotein enzyme telomerase adds telomeric DNA onto chromosome ends and is normally regulated so that telomeric DNA lengths are kept within defined bounds. In the telomerase RNA gene from the yeast *Kluyveromyces lactis*, specific mutations that alter telomeric DNA sequences result in telomeres elongating to up to 100 times their normal length and impair cell growth. Some mutations cause immediate elongation whereas others behave like genetic time bombs, causing elongation only after a latent period of hundreds of generations.

DNA ends cannot be fully replicated by normal DNA polymerases and the ends of linear chromosomes, the telomeres, would gradually shrink in the absence of a specialized means of maintaining them (reviewed in ref. 1). The telomeric DNA of almost all eukaryotic cells is composed of arrays of tandem copies of a short DNA repeat and is kept from progressively shortening by telomerase, a ribonucleoprotein enzyme that adds telomeric repeats onto telomeric DNA termini¹⁻³. This specialized cellular reverse transcriptase uses part of its RNA moiety as the template for the synthesis of one strand of the telomeric repeat DNA at the chromosomal ends^{4,5}.

Although the enzymatic properties of telomerase have been studied *in vitro*^{2,6-9}, little is known about regulation of its activity *in vivo*. At one extreme, telomerase activity appears to be absent in many human somatic cells and telomeres shorten during an individual's lifetime^{10,11}. It has been proposed that the finite cell division capacity of human somatic cells is limited by telomere length¹⁰. This is consistent with reports that telomerase activity is often high in cancers and immortalized tissue culture cells¹²⁻¹⁶. Human cells expressing telomerase as well as cells of lower eukaryotes have tracts of telomeric repeats that do not continually shrink. Instead, their telomeric arrays are maintained within a length distribution that varies widely between species but is generally constant and often quite narrow within a species^{17,18}.

Previous analysis of individual telomeres in clonal lines of the yeast *Saccharomyces cerevisiae* showed that telomeres could lengthen and shorten in small increments in a seemingly stochastic manner, most likely through telomerase action and incomplete replication respectively¹⁹. However, this length heterogeneity was confined within clear upper and lower limits¹⁹. High-copy-number circular plasmids carrying arrays of telomeric repeats can cause modest telomere lengthening in *S. cerevisiae*, consistent with their competing for a negative regulator of telomere length²⁰. This negative regulator might be the Rap1 protein, which binds to *S. cerevisiae* telomeric repeats^{21,22} and can cause moderate degrees of telomere elongation when mutationally altered²³.

Here we report the identification of the telomerase RNA gene from the budding yeast *Kluyveromyces lactis*, a species with unusually long (25 base pairs) telomeric repeat units, whose total tract length is normally tightly controlled between 0.25 and 0.5 kilobases²⁴. We show that, as expected for a telomerase RNA gene, mutations within the template region are incorporated into the telomeric repeats of *K. lactis* cells. Initially, the mutant repeats were incorporated only at the most terminal region of the repeat array, but over many cell divisions they gradually spread further in from the terminus. Unexpectedly, some template mutations led to runaway telomere elongation, which began immediately in one mutant but whose onset was extremely

delayed in other mutants. We provide evidence that the long-delayed lengthening occurs after a critical number of the internal wild type telomeric repeats have been replaced by mutant repeats. Hence, telomere length is under active negative regulation, and the internal telomeric repeats play an essential role in controlling telomerase action at chromosomal termini.

Cloning the telomerase RNA gene

In several budding yeast species the telomeric repeat units are much longer than the 5–8-base-pair (bp) repeats typical of most eukaryotes^{24,25}. These yeasts include *K. lactis* and *Candida albicans*, with 25-bp and 23-bp repeat units respectively. These long telomeric repeats are also highly unusual in not being G-rich overall, which raised the questions of whether they function and are maintained by mechanisms similar to those of other organisms, and in particular, whether they are synthesized by telomerase. We took advantage of the large size of the *K. lactis* telomeric repeat unit by directly using its sequence as a hybridization probe to screen for candidate telomerase RNA genes in the genome of haploid *K. lactis*. By this method we cloned a genomic region located at a non-telomeric position and containing a 30-bp sequence consisting of one and a fifth *K. lactis* telomeric DNA repeats (Fig. 1a). This telomeric homology was present in a single copy genomic sequence (which we call *TER1*) encoding a ~1.3-kilobase (kb) non-polyadenylated RNA molecule of the expected strand to function as a telomerase RNA (Fig. 1b).

To determine whether the telomere homology sequence was the template sequence of a telomerase RNA gene, base substitutions were made within it (Fig. 1c). We initially replaced the wild-type *TER1* sequence in haploid *K. lactis* with a *TER1* sequence carrying a pair of base changes within the putative template that generated a *Bgl*II restriction site (*ter1-Bgl*II; Fig. 1c). This *Bgl*II mutation resulted in the appearance of the corresponding sequence change in the telomeric repeats, as demonstrated by the following criteria: hybridization using a ³²P-labelled oligonucleotide probe specific for the expected mutant sequence (Fig. 2b), the appearance of the predicted restriction site in terminal regions of telomeres (Fig. 2a, b), and sequencing of a cloned telomere from a transformant (data not shown). The mutant *Bgl*II repeats were initially confined to the very termini of the telomeric tracts, as seen in Fig. 2a and b, which shows Southern blotting analyses of the telomeric DNA from cells transformed by the *ter1-Bgl*II sequence. *Bgl*II digestion completely removed the mutant telomeric repeats (Fig. 2b), but by the second passage only shortened the telomeres by ~100 bp or less, with no appreciable reduction in hybridization to the wild-type telomeric probe (Fig. 2a).

Wild-type repeats continued to comprise the more internal part of each telomeric repeat array (Fig. 2a). The relative num-

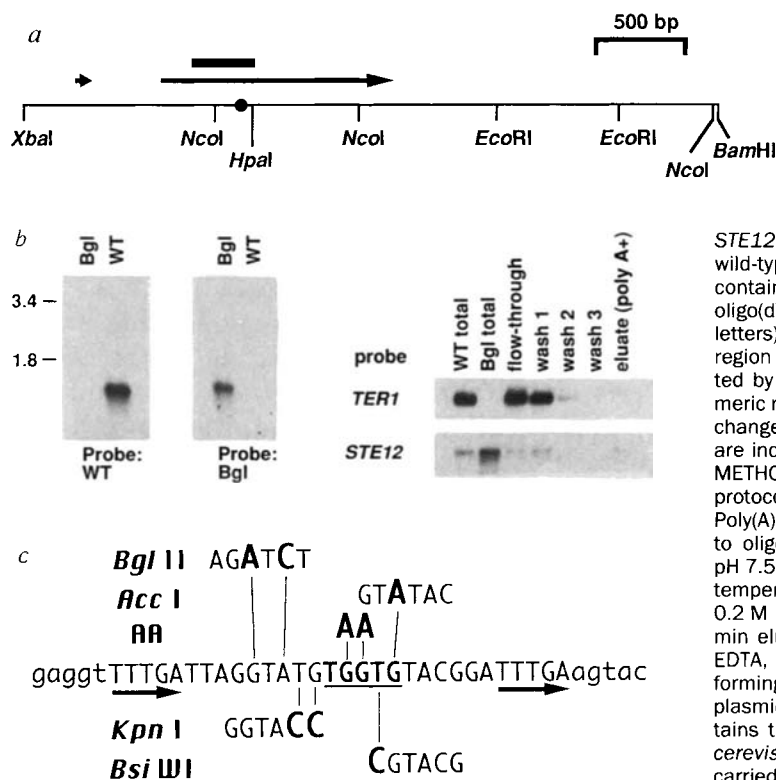


FIG. 1 The *K. lactis* telomerase RNA gene (*TER1*). a, Genomic map of the *TER1* gene region. The long arrow indicates the position of the *TER1* gene, pointing in the direction of transcription; black dot, position of the template sequence; bar, approximate extent of the *ter1*- $\Delta 7$ allele used in this work. A *trnA^{le}* gene (short arrow) that is transcribed in the same direction as *TER1* and positions of restriction sites are also shown. b, Identification of the *TER1* RNA. Left, northern blot with RNA samples

bers of wild-type and mutant repeats were estimated directly by comparing the ratio of signal seen with telomeric fragments to that seen with single-copy telomerase RNA gene fragments (marked with arrows). For each telomere, this estimate of the number of *Bgl*II repeats was consistent with the extent of shortening caused by the secondary *Bgl*II digestion. By the second rich-medium passage (60–75 generations), after replacing wild-type *TER1* with mutated allele *ter1*-*Bgl*II, telomeres in the *ter1*-*Bgl*II cells contained 1–5 mutant repeats. Similar results were obtained for another mutated *ter1*, *ter1*-AA, with the GG to AA mutation shown in Fig. 1c. Again, mutant repeats were incorporated (Fig. 2d), and the hybridization signals to an oligonucleotide probe specific to wild-type telomeric repeats were roughly comparable in transformed and untransformed cells.

We also constructed strains carrying a deletion of the *TER1* gene sequence. Deleting a ~300-bp region that includes the template of the *TER1* gene and inserting the *S. cerevisiae* *URA3* gene in its place (the *ter1*- $\Delta 7$ allele; Fig. 1a) led to a gradual shortening of all telomeres, at a rate of ~5 bp per cell division (Fig. 2e). This steady loss of telomeric sequences was accompanied by a gradual decline in cell growth rates, and within ~100 generations the majority of cells were inviable. These results further confirmed the role of the *K. lactis* *TER1* gene in telomere maintenance, and are similar to those seen with *est1*⁻ and *tlc1*⁻ gene deletions in *S. cerevisiae*^{26,27}. As with *est1*⁻ cells of *S. cerevisiae*²⁸, some *K. lactis* cells carrying the *ter1*- $\Delta 7$ allele survived the initial senescence event. The behaviour of these post-senescence survivors will be described elsewhere.

These results demonstrated that the chromosome-internal 30-bp telomere-homology region functions as the template for the synthesis of telomeric repeats. We therefore conclude that we

prepared from either *K. lactis* wild-type cells (7B520) or from cells carrying only *ter1*-*Bgl*II (strain 28-14) and hybridized with either a probe specific to wild-type *TER1* template (KLAC1-25 (ACGGATTGATTAGGTATGGTGTG) at 49 °C) or *ter1*-*Bgl*II mutant (KLTELBGL (GATTAGTCTGTGG) at 40 °C). Positions and sizes (in kb) of rRNA markers are given on the left. Right, the apparent lack of polyadenylation of *TER1* RNA is shown on a northern blot probed with either wild-type *TER1* template sequence or

STE12 gene as control. The first two lanes contain total RNA from either wild-type cells or strain 28-14 cells (containing *ter1*-*Bgl*II). Other lanes contain flow-through, washes or eluate from wild-type RNA bound to oligo(dT) beads. c, The 30-bp region of telomere homology (upper-case letters) of the *TER1* gene. The strand shown is the complement of a region of RNA sequence. The duplicated part of the sequence is indicated by arrows and the conserved-sequence motif present in the telomeric repeats of several yeast species²⁴ is underlined. Positions of base changes made to create various mutations within the template region are indicated with the names of the alleles created shown on the left. METHODS. RNA isolation³⁹ and northern analysis⁴⁰ followed standard protocols. Hybridization was done as described for DNA analysis. Poly(A)⁺ RNA was isolated by binding total *K. lactis* RNA (0.06 mg ml⁻¹) to oligo(dT) cellulose (10 mg ml⁻¹) in 8 ml 0.5 M NaCl, 10 mM Tris, pH 7.5, 1 mM EDTA, 0.2% SDS, with gentle shaking for 1 hour at room temperature. Three 5-min washes of 5 ml each were made with 5 ml 0.2 M NaCl, 10 mM Tris, pH 7.5, 1 mM EDTA, 0.2% SDS. Three 5-min elutions were then made with 1.2 ml 10 mM Tris, pH 7.5, 1 mM EDTA, 0.2% SDS. In c, each of the mutations was made by first transforming the haploid *K. lactis* strain 7B520 (ref. 41) with the integrative plasmid pTER-BX-UA carrying a template mutation. pTER-BX-UA contains the *TER1* gene as an ~4-kb *Bam*HI-*Xba*I fragment and the *S. cerevisiae* *URA3* gene as a selectable marker. *Ura3*⁺ transformants carried the plasmid integrated by homologous recombination at the native *TER1* locus and carried two copies of *TER1*, usually one mutant and one wild type. Strains carrying only a single mutant copy of *TER1* were then created by selection for *Ura3*⁻ cells that had undergone plasmid 'pop-out', on medium containing 5-fluoro-orotic acid⁴² and screening for the presence of the *TER1* template mutation by Southern hybridization. The nucleotide sequence data reported in this paper will appear in the EMBL, Genbank and DDBJ nucleotide sequence databases under the accession number U31465.

have identified the telomerase RNA gene of *K. lactis* (*TER1*, as it is the functional homologue of the telomerase RNA gene *TER* of the ciliate *Tetrahymena*²⁹). A map of the cloned genomic DNA region containing the *K. lactis* *TER1* gene is shown in Fig. 1a. The overall size of the *TER1* RNA and the relative position of the template region are similar to those of the telomerase RNA gene *TLC1* from *S. cerevisiae*²⁷. However, there is no appreciable sequence homology between the two yeast RNA coding sequences.

Effects of mutating a conserved motif

The long telomeric repeats in *K. lactis* share a short sequence motif, TG_{2,3}TG, with both the long and short telomeric repeat sequences in other budding yeasts, including *S. cerevisiae*²⁴. Although the wild-type *K. lactis* telomeric repeat is not G-rich overall, this conserved yeast motif has the two to five consecutive G residues found in virtually all known telomeric repeats. It has been proposed that this conservation of G residues is related to the *in vitro* capability of the G-rich telomeric sequences of various eukaryotes to form unusual DNA structures³⁰ such as G-quartets^{30,33}. Telomerase template mutations that alter these G runs produce severe cellular phenotypes in *Tetrahymena thermophila*²⁹. But surprisingly, when the wild-type *TER1* gene was replaced by a mutant telomerase RNA gene in which the pair of consecutive G residues within the *K. lactis* conserved motif were changed to A residues (the *ter1*-AA gene; Fig. 1c), no deleterious growth phenotype was produced in the mitotically dividing cells (data not shown) and there was no alteration in telomere length over at least 600 cell divisions (Fig. 2c, d, and data not shown).

Two other mutations within the conserved motif of the *TER1* gene template sequence were also created: a T-to-C change (*ter1*-

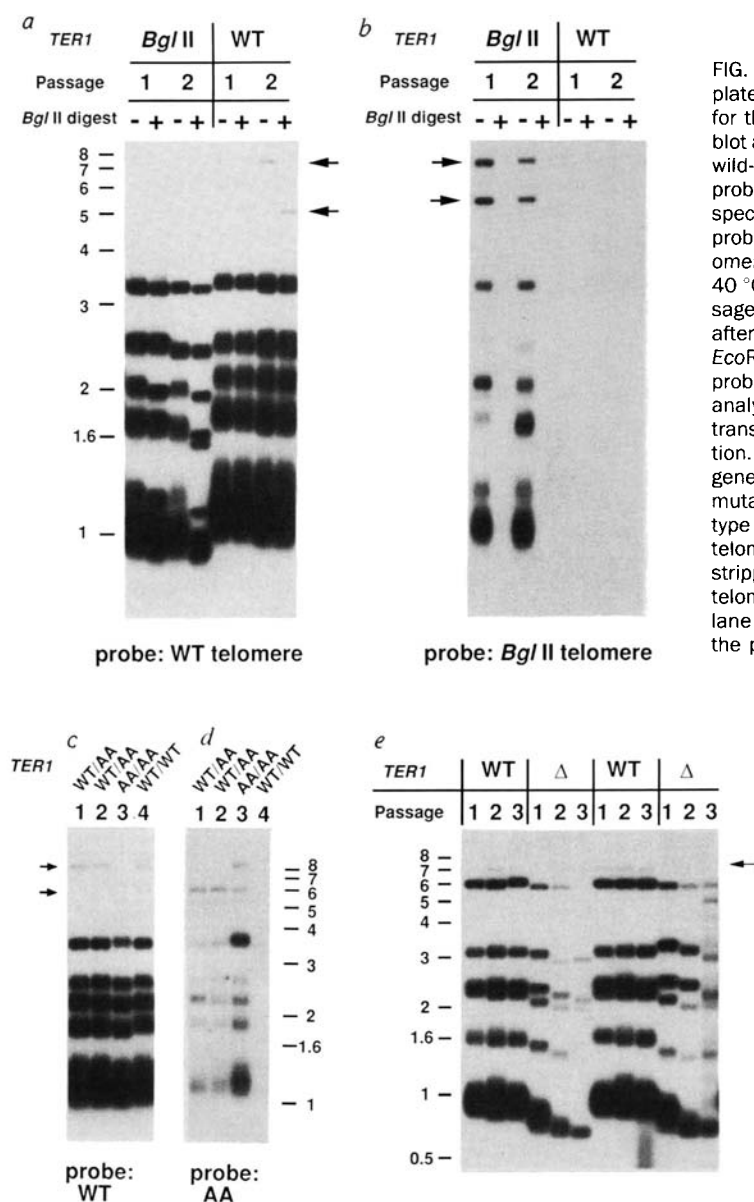


FIG. 2 The telomere homology sequence in *TER1* functions as a template for the synthesis of *K. lactis* telomeric repeats and is essential for the normal maintenance of *K. lactis* telomeres. **a** and **b**, Southern blot analysis of DNAs from wild-type *K. lactis* and a *K. lactis* strain whose wild-type *TER1* allele had been replaced by the *ter1-BglII* mutation. The probe used in **a** was an oligonucleotide (KLAC1-25; hybridized at 55 °C) specific for wild-type telomeric repeats; **b** shows the same filter after probing with an oligonucleotide, specific for the *ter1-BglII* mutant telomeric repeats, that spans the *BglII* site (KLTEL BGL; hybridized at 40 °C). Two consecutive passages of both strains are shown. The passages of the *ter1-BglII* mutant strain (clone 28-14) are the first two after the creation of the strain. Each DNA sample was digested with *EcoRI* or *EcoRI* + *BglII* as indicated. After *BglII* digestion the KLTEL BGL probe can no longer hybridize to *BglII* repeats. **c** and **d**, Southern blot analysis of *EcoRI*-digested DNAs from *K. lactis* strains that had been transformed with an integrative plasmid containing the *ter1-AA* mutation. Two transformants (lanes 1, 2) contain one copy of wild-type *TER1* gene and one copy of *ter1-AA*. The transformant in lane 3 carries two mutant *ter1-AA* alleles and the transformant in lane 4 carries two wild-type copies of *TER1*. **c**, Hybridization to a probe specific to wild-type telomeres (KLTELWT (GATTAGGTATGTGG); 40 °C); **d**, same filter after stripping and rehybridization with a probe specific to the AA form of telomeres (KLAA (GTATGTAATGTACG); 40 °C). The weak hybridization in lane 4 is due to nonspecific hybridization and residual signal from the prior KLTELWT hybridization. **e**, Progressive loss of telomeric DNA in *K. lactis* clones containing a deletion of the *TER1* template region. Shown are two wild-type *K. lactis* Y-160 (gift from I. Herskowitz) clones and two independent *ter1Δ7* mutants of Y-160 digested with *EcoRI* and probed with an oligonucleotide specific to *K. lactis* telomeric repeats (KLAC 1-25; 50 °C). DNA samples prepared from the original transformation colony and two serial restreaks after the initial isolation of the mutants strains are shown as passages 1–3. Arrows indicate positions of fragments containing the *TER1* gene. DNA size markers are indicated in kb. *K. lactis* DNA was prepared from cells grown at 30 °C in liquid YPD medium. Southern blots were prepared using Hybond N⁺ filters (Amersham) according to the method recommended by the manufacturer, and DNA was crosslinked to the membrane using a Stratalinker 1800 (Stratagene). Hybridizations were carried out according to ref. 43 using ³²P-labeled oligonucleotide probes. *K. lactis* cells were passaged by restreaking of colonies twice a week down to single cells onto YPD-rich medium with growth at 30 °C. Each passage represents 20–25 cell divisions. In certain cases, 20–25 cell divisions on a selective medium plate preceded the first YPD passage.

BsiWI) and a G-to-A change (*ter1-AccI*; each is named for the restriction enzyme site created by the mutation; Fig. 1a). Haploid strains that were hetero-allelic at the *TER1* chromosomal region for *ter1-AccI* plus a copy of wild-type *TER1*, or for *ter1-BsiWI* and wild-type *TER1*, showed only slight telomere lengthening (Fig. 3a). In contrast, after selection for loss of the wild-type *TER1* gene (Fig. 1 legend), so that only the *ter1-AccI* or *ter1-BsiWI* gene was present, the telomeres became markedly elongated. This loss of telomere length control was particularly striking in the *ter1-AccI* mutant: by ~50 cell divisions after the loss of the wild-type *TER1* allele (the first passage) most of the telomeric hybridization was either at limit mobility (>10 kb) or did not enter the agarose gel (Fig. 3a). This massive lengthening was also reflected by greatly increased hybridization to an oligonucleotide probe which detects both wild-type and mutant *K. lactis* telomeric repeats. All the length increase was attributable to addition of *AccI* mutant telomeric repeats, because digestion with *AccI* reduced the telomeric restriction fragments to almost wild-type length, releasing large amounts of small *AccI* telomeric repeat fragments (Fig. 3a, and data not shown). The extreme lengthening of *ter1-AccI* mutants appeared to be complete within 75 cell divisions and was invariably accompanied by poor cell

growth, characterized by rough, variably sized colonies and a high percentage of inviable cells.

Telomere lengthening was less severe in the *BsiWI* mutant strain. After the first passage of cells carrying only the *ter1-BsiWI* allele, the telomeres were up to 1.5 kb longer and considerably more heterogeneous than wild-type telomeres, becoming even longer by the second passage (Fig. 3a). Again, digestion with *BsiWI* removed the mutant telomeric repeats and reduced the elongated telomeric fragments to approximately wild-type length (Fig. 3a). With additional passaging, the telomeres of *ter1-BsiWI* strains continued to gradually elongate (data not shown).

Delayed runaway telomere phenotype

A strikingly different manner of telomere elongation was displayed by two other *K. lactis* *TER1* mutations, *ter1-BglII* and *ter1-KpnI* (Fig. 1b). These mutations, unlike the *AccI* and *BsiWI* mutations, required elimination of most of the pre-existing wild-type telomeric repeats before they produced telomere elongation. An example of this is shown in Fig. 3b, which shows telomeric DNA of cells carrying only the *ter1-BglII* allele monitored for several hundred cell divisions. Initially (Figs 2a and 3b, passages

2–20) this mutation produced telomeres that were approximately two thirds of normal length. During this period, *Bgl*II mutant telomeric repeats slowly replaced wild-type repeats at progressively more internal positions. The presence of a restriction site in the telomeric repeat unit allowed easy measurement of the penetration of mutant repeats into more internal areas of telo-

meric repeat arrays. This was shown by the progressive decrease in the length of the telomeric fragments remaining after double digestion with *Bgl*II+*Xba*I relative to *Xba*I digestion alone (Fig. 3b). Gradual turnover of wild-type telomeric repeats is the expected outcome of the combined actions of telomere shortening and telomerase. As individual telomeres occasionally shrink

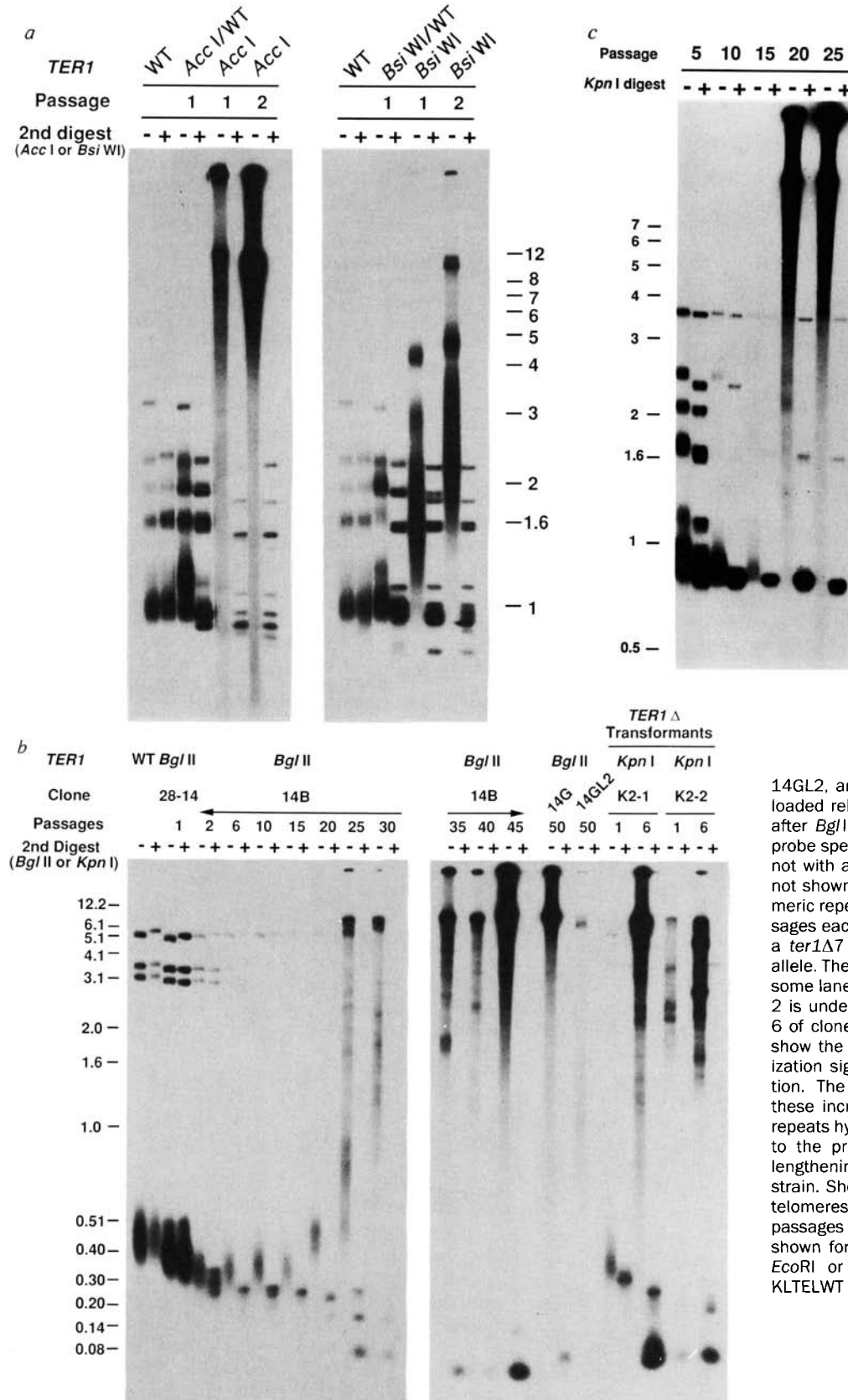


FIG. 3 Telomere-lengthening mutants. **a**, Immediate telomere lengthening produced by two *TER1* template mutations. DNA samples from *K. lactis* cells containing *ter1*-*AccI* or *ter1*-*BsiWI*, either with or without a second, wild-type allele of *TER1* were digested with *EcoRI* or *EcoRI* + *AccI* or *EcoRI* + *BsiWI*. The Southern blot was probed with a *K. lactis* telomeric sequence (KLAC1-25 at 50 °C). DNA prepared from the initial passage of the plasmid-integration strains and the first two rich-medium passages of the mutant 'pop-out' strains are shown. Positions of DNA size markers in kb are indicated. **b**, Highly delayed telomere lengthening produced by *TER1* template mutations. The *ter1*-*Bgl*II *K. lactis* strain designated 14B was followed over a period of 45 passages (each passage being a restreak to single cells and constituting 20–25 cell divisions) after isolation from its parental strain 28-14. The first passage of clone 28-14 (containing two copies of the *ter1*-*Bgl*II allele) is also shown. Genomic DNA was digested with either *Xba*I or *Xba*I + *Bgl*II and the Southern blot was probed with a *K. lactis* telomeric sequence (KLAC1-25 at 50 °C). DNA from the 50th passage of *ter1*-*Bgl*II clone 14G and its faster-growing derivative

14GL2, are also shown (14GL2 is slightly overloaded relative to 14G). Small fragments visible after *Bgl*II digestion were also detected with a probe specific for wild-type telomeric repeats but not with a probe specific for *Bgl*II repeats (data not shown) and therefore contain wild-type telomeric repeats. The four right lanes show two passages each of two independent transformants of a *ter1*Δ7 strain that had received a *ter1*-*KpnI* allele. There is variability in DNA loading between some lanes (for example, passage 1 of clone K2-2 is underloaded). Loadings of passages 1 and 6 of clone K2-1 were well matched and clearly show the massive increase in telomeric hybridization signal associated with telomere elongation. The signals in this blot underrepresent these increases, as the *Bgl*II and *KpnI* mutant repeats hybridize less well than wild-type repeats to the probe used here. **c**, Delayed telomere lengthening phenotype in a *K. lactis* *ter1*-*KpnI* strain. Shown is a Southern blot analysis of the telomeres in a *ter1*-*KpnI* strain followed over 25 passages (500–625 generations). The two lanes shown for each sample are digests with either *EcoRI* or *EcoRI* + *KpnI*. The probe used was KLTELWT at 40 °C.

down to the minimal length of their permitted size range, only the mutant *Bgl*II telomerase is present to elongate them. Telomerase mediated repeat turnover does not, however, appear to be the sole source of change occurring near telomeres in *ter1-Bgl*II strains. The relatively frequent loss or alteration in size of telomeric bands in *ter1-Bgl*II but not wild-type strains (data not shown, and Fig. 3b: note loss of two large telomeric fragments in clone 14B) suggests that gene conversion or recombination events occur frequently near the telomeres of *ter1-Bgl*II strains.

Despite the gradual replacement of wild-type repeats with *Bgl*II repeats, during the first 15 passages (300–375 cell divisions) total telomeric lengths in the *ter1-Bgl*II clone 14B remained the same (Fig. 3b). However, after 20 passages (400–500 cell divisions) the telomeres in this lineage had elongated slightly, and after 25 passages had begun to increase sharply in size, eventually reaching a plateau averaging up to several tens of kilobases, or ~100 times normal length, as estimated by quantification of hybridization signals. This delayed telomere lengthening was accompanied by the onset of poor cell growth, a rough colony morphology, and decreased cell viability, phenotypes similar to those of the *ter1-AccI* mutant. A partial telomeric fragment cloned from a *ter1-Bgl*II strain after elongation had occurred showed that the telomeric repeats had not acquired any additional sequence alterations. As controls, three different lineages of wild type *K. lactis* cells were grown for 60 passages (~1,200–1,500 cell divisions each), and none showed telomeric size changes or produced small, rough-looking colonies. We therefore conclude that the delayed runaway lengthening phenomenon is characteristic of the *Bgl*II mutation and defines a novel telomere phenotype.

Five of nine independent lineages of *ter1-Bgl*II mutant cells started runaway lengthening after 20–25 passages (400–625 cell divisions). In some of these lineages the extreme lengthening began and was complete within five or fewer passages (100–125 cell divisions). In other lineages, such as clone 14B of Fig. 3b, a similar degree of lengthening required considerably more passages. In those lineages that did not display runaway elongation, the telomeres typically were lengthened up to approximately wild-type length and displayed a faint smear of telomeric hybridization extending up to limit mobility. We believe that this smear represents loss of telomere-length regulation in a subset of the cell population at the time of analysis. Consistent with this is the observation that all nine *ter1-Bgl*II lineages produced occasional small rough colonies, which when analysed by Southern analysis, were found to contain cells with extremely long telomeres.

Although not all *ter1-Bgl*II cell lineages passaged 20 or more times displayed runaway telomere elongation, the occurrence of runaway telomere elongation correlated with the size of *Xba*I + *Bgl*II fragments that hybridized to wild-type telomeric repeats. In lineages that had undergone runaway elongation, these *Xba*I + *Bgl*II fragments were mostly shorter than ~250 bp, whereas in lineages that still had short telomeres, these fragments were mostly larger than ~250 bp. Hence, runaway elongation may depend upon a sufficient degree of loss of the residual wild-type telomeric repeats. Several observations suggest that this replacement of the most internally located wild-type repeats may involve processes other than telomerase-mediated repeat addition. First, following the steady progressive encroachment of *Bgl*II repeats during the initial ~15 passages, further encroachment occurred in jumps and was highly variable between clonal lineages (Fig. 3b, and data not shown). Second, in the *ter1-Bgl*II lineages that had undergone elongation, many of the small *Xba*I + *Bgl*II fragments containing wild-type repeats were too small (<100 bp) to contain the expected ~120 bp of subtelomeric sequences, and therefore had been altered in some additional way besides telomerase-mediated repeat turnover (Fig. 3b, and data not shown). Together with the observation that subtelomeric gene conversion events were frequent in *ter1-Bgl*II strains (Fig. 3) these data suggest that sporadic recombinational

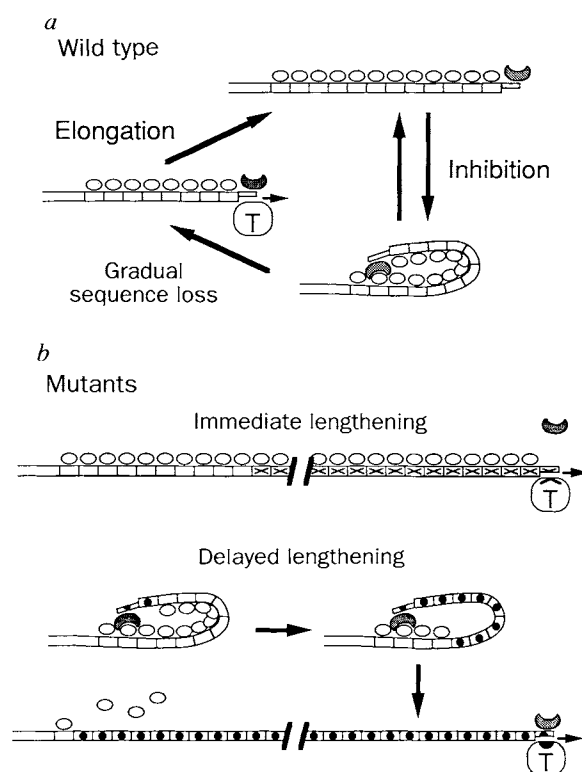


FIG. 4 Telomere-length regulation. *a*, Model for the regulation of telomere length in wild-type *K. lactis*. The ability of telomerase to add new telomeric repeats onto chromosomal termini is proposed to be negatively regulated by one or more protein factors that bind to the telomeric repeats. Elongation of a telomere by telomerase will increase the number of molecules of an inhibitory double-strand telomere-binding protein (DS-TBP; shown as ovals) in the vicinity of that telomere and favour inhibition of further elongation, by favouring interaction with a terminus-specific binding protein (shaded half moon shape), or by other means (see text). Progressive sequence loss resulting from DNA replication without telomerase addition would shorten the telomere until elongation becomes favoured. *b*, Immediate and delayed telomere-lengthening phenotype caused by telomerase template mutations. For the immediate telomere-lengthening phenotype, part of the regulatory system at the very terminus is altered by the telomeric DNA sequence change, perhaps by preventing the binding of a terminus-specific binding protein. For the highly delayed telomere-lengthening phenotype, the telomeric DNA-sequence change prevents or alters the binding of a DS-TBP. The elongation phenotype is not manifested until the majority of the internal wild-type repeats are turned over through the action of telomerase, and perhaps recombination and gene conversion (see text).

events that eliminate some of the most internal wild type repeats trigger runaway telomere lengthening in *ter1-Bgl*II strains.

The *ter1-Bgl*II lines that had undergone runaway telomere lengthening and were growing as small, rough-looking colonies occasionally produced larger, smoother 'revertant' colonies. These faster-growing derivatives retained the *Bgl*II site in the *ter1* template region and displayed variable degrees of telomere reduction (for example, note the large decrease in hybridization in clone 14GL2 in Fig. 3b). This situation is reminiscent of the occasional abrupt shortening of the moderately long telomeres produced by *RAP1* mutation in *S. cerevisiae*²³. This observation, together with the poor growth of both the immediate and the delayed lengthening mutants, suggests that extremely long telomeres may inhibit growth of *K. lactis* cells. The relatively common occurrence (estimated frequency ~10⁻³) of the 'revertants', of the long telomere form of *ter1-Bgl*II mutants indicates that *K. lactis* is able to eliminate excess telomeric sequences, perhaps through recombination or DNA breakage. The overall plateau

in the amounts of telomeric DNA in the *K. lactis* *AccI* and *Bgl*II telomere runaway mutants may therefore reflect a balance between continuous uncontrolled telomerase activity and processes that occasionally act to shorten telomeres drastically, combined with continuous selection for cells lacking lethally elongated telomeres.

Model for telomere-length regulation

The *TER1* mutations simultaneously alter both the telomerase enzyme and the telomeric DNA. Thus in principle, elongation could result from a mutation increasing the intrinsic activity of the telomerase enzyme. This is unlikely for the *ter1-AccI* and *ter1-Bsi*WI mutations as their phenotypes are largely recessive to wild-type *TER1* (Fig. 3a). This explanation also does not readily account for the great delay before the onset of lengthening in the *ter1-Bgl*II and *ter1-KpnI* mutants. Instead, the concomitant onset of the delayed runaway telomere phenotype of the *ter1-Bgl*II mutant with loss of most of the internal wild type telomeric repeats suggests that the mutant *Bgl*II telomeric repeats are defective in negative regulation of telomere length. We propose that the latent period before length regulation is lost corresponds to the period in which the internal wild-type repeats are progressively replaced by *Bgl*II repeats (see *ter1-Bgl*II clone 14B in Fig. 3b). Eventually the number of internal wild-type repeats falls below a critical threshold level required for length regulation. We propose at this point that the negative control over length regulation is lost, allowing unhindered telomeric lengthening by telomerase. Hence, in this model, the delay in the manifestation of the telomere lengthening phenotype is caused by the slow rate (~1 bp per generation) of replacement of internal wild-type repeats by mutant *Bgl*II repeats. This hypothesis predicts that introducing a delayed lengthening mutant *TER1* allele into a *ter1-Δ7* strain in which cells have already undergone progressive telomere shortening (Fig. 2e) will eliminate or shorten the latent period preceding the onset of runaway telomere lengthening. To test this, cells that were carrying only the deleted *ter1-Δ7* allele and were in the process of gradual telomere shortening were transformed with either the *ter1-Bgl*II or the *ter1-KpnI* gene. The *ter1-KpnI* mutant gene (Fig. 1c) causes an initial short telomere phenotype and a delayed lengthening phenotype very similar to those of *ter1-Bgl*II mutants (Fig. 3c). Two transformants were recovered, both carrying an integrated *ter1-KpnI* gene. One (clone K2-1) initially displayed the shorter-than-normal telomeres characteristic of the *ter1-KpnI* early phenotype (Fig. 3c). However, within the next five passages its telomeres became greatly elongated. In the other independent *ter1-KpnI* transformant (clone K2-2), the telomeres were already highly elongated as soon as it was analysed (Fig. 3b). Based on the sizes of the telomeric fragments remaining after *XbaI* + *KpnI* digestion (Fig. 3b), telomeres from the first passage of clone K2-1 contained more wild-type repeats than either the same clone at passage six, or telomeres of clone K2-2 at either passage. These data support the hypothesis that the eventual replacement of the wild-type internal repeats causes the delayed telomere-lengthening phenotypes of the original *ter1-Bgl*II and *ter1-KpnI* mutants.

Similar to our results with *ter1-KpnI* cells, we have shown that transformation of other template mutations into *ter1-Δ7 K. lactis* cells results in greater degrees of telomere length disruption than is otherwise produced by the mutations. Under these circumstances, the *ter1-Bsi*WI allele produces runaway telomere lengthening and poor cell growth, and the *ter1-AA* allele produces at least moderate telomere lengthening (data not shown). These results further emphasize the importance of remaining wild-type telomeric repeats in telomere length regulation.

We propose a model for telomere-length regulation based on these novel findings. In this model, interaction of a sequence-specific double-stranded telomeric DNA-binding protein (DS-TBP) with the tract of duplex telomeric repeats directly or indirectly inhibits telomerase-mediated lengthening of that telomere (Fig. 4). In wild-type cells, any telomeres that become relatively long will have more molecules of the DS-TBP near their termini than shorter telomeres, and thus will be more likely to be in a state refractory to telomerase action. Conversely, shorter telomeres will be more accessible to telomerase action than longer ones. This would provide the basis for steady-state regulation of telomere length operating at the level of the telomere DNA-protein complex itself.

By this model, mutated telomeric repeats could potentially interfere with telomere-length regulation in at least two ways, giving rise to immediate or delayed effects. We propose that the *Bgl*II and *KpnI* mutations disrupt or alter binding of the DS-TBP to telomeric repeats. Hence, once the tract length of the remaining wild-type internal repeats becomes too short, unregulated telomere lengthening occurs. Once this elongation is underway, the telomeric terminus will be moved progressively further away from any remaining wild-type repeats, which could further reduce the normal regulatory interaction between the DS-TBP and the terminus and thereby favour runaway telomeric growth. These delayed-lengthening telomeric mutations could inhibit DS-TBP binding or alter it in a way that prevents formation of a higher-order telomeric protein-DNA complex.

We propose that the immediate telomere lengthening in the *ter1-AccI* and the *ter1-Bsi*WI mutants is caused by disrupting a structure at the very terminus of the chromosome, making it unresponsive to the inhibition normally mediated by the DS-TBP. Two of the possible mechanisms for this immediate effect are a more severe binding disruption of the DS-TBP postulated to be involved in the delayed lengthening effect, and/or disrupted binding of a protein that specifically recognizes telomeric termini, such as those that bind the 3' single-strand telomeric overhang in ciliated protozoans^{34,35}.

In *S. cerevisiae*, the Rap1 protein that binds double-stranded telomeric DNA influences telomere-length regulation^{23,36}. Therefore the *K. lactis* Rap1p homologue may be the DS-TBP in our model³⁷. Rap1 is an attractive candidate because *S. cerevisiae* Rap1 is known to participate in one type of telomere length-dependent repressive effect mediated 'in cis': transcriptional silencing of genes placed near telomeres³⁸. We are currently investigating *K. lactis* telomere-binding proteins and testing directly whether the *TER1* mutations we have analysed here affect their binding *in vitro*.

The runaway telomere mutants obtained by altering telomeric DNA sequences have shown that negative telomere-length regulation is associated with optimal cell viability. This work has shown that a critical role in this regulation is played by the composition of the telomeric DNA itself. These mutants also reveal a novel way in which an essential cellular reverse transcriptase can escape from its normal mode of control and be converted into a detrimental overproducer of its DNA product. Perhaps such escape from control of synthesis of its DNA product may have been one step by which a cellular reverse transcriptase could have evolved into a viral reverse transcriptase. □

Received 12 April; accepted 28 June 1995.

- Blackburn, E. H. *Cell* **77**, 621–623 (1994).
- Blackburn, E. H. *A. Rev. Biochem.* **61**, 113–129 (1992).
- Pardue, M. L. *Curr. Opin. Genet. Dev.* **4**, 845–850 (1994).
- Greider, C. W. & Blackburn, E. H. *Nature* **337**, 331–337 (1989).
- Shippen, L. D. & Blackburn, E. H. *Science* **247**, 546–552 (1990).
- Greider, C. W. *Molec. cell. Biology* **11**, 4572–4580 (1991).
- Harrington, L. A. & Greider, C. W. *Nature* **353**, 451–454 (1991).
- Lee, M. & Blackburn, E. H. *Molec. cell Biol.* **13**, 6586–6599 (1993).
- Shippen, D. E., Blackburn, E. H. & Price, C. M. *Proc. natn. Acad. Sci. U.S.A.* **91**, 405–409 (1994).
- Harley, C. B., Futcher, A. B. & Greider, C. W. *Nature* **345**, 458–460 (1990).
- Hastie, N. D. *et al. Nature* **346**, 866–868 (1990).
- Morin, G. B. *Cell* **59**, 521–529 (1989).
- Prowse, K. R., Avilion, A. A. & Greider, C. W. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1493–1497 (1993).
- Shay, J. W., Wright, W. E., Brasiskyte, D. & van der Hagen, B. A. *Oncogene* **8**, 1407–1413 (1993).
- Counter, C. M., Hirt, H. W., Bacchetti, S. & Harley, C. B. *Proc. natn. Acad. Sci. U.S.A.* **91**, 2900–2904 (1994).

16. Klingerhutz, A. J., Barber, S., Smith, P. P., Dyer, K. & McDougall, J. K. *Molec. cell. Biol.* **14**, 961–969 (1994).
17. Blackburn, E. H. *Nature* **350**, 569–573 (1991).
18. Zakian, V. A. *Rev. Genet.* **23**, 579–604 (1989).
19. Shampay, J. & Blackburn, E. H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 534–538 (1988).
20. Runge, K. W. & Zakian, V. A. *Molec. cell. Biol.* **9**, 1488–1497 (1989).
21. Berman, J., Tachibana, C. Y. & Tye, B.-K. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3713–3717 (1986).
22. Conrad, M. N., Wright, J. H., Wolf, A. J. & Zakian, V. A. *Cell* **63**, 739–750 (1990).
23. Kyron, G., Boakye, K. A. & Lustig, A. J. *Molec. cell. Biol.* **12**, 5159–5173 (1992).
24. McEachern, M. J. & Blackburn, E. H. *Proc. natn. Acad. Sci. U.S.A.* **91**, 3453–3457 (1994).
25. McEachern, M. J. & Hicks, J. B. *Molec. cell. Biol.* **13**, 551–560 (1993).
26. Lundblad, V. & Szostak, J. W. *Cell* **57**, 633–643 (1989).
27. Singer, M. S. & Gottschling, D. E. *Science* **266**, 404–409 (1994).
28. Lundblad, V. & Blackburn, E. H. *Cell* **73**, 347–360 (1993).
29. Yu, G.-L., Bradley, J. D., Attardi, L. D. & Blackburn, E. H. *Nature* **344**, 126–132 (1990).
30. Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I. J. & Blackburn, E. H. *Cell* **51**, 899–908 (1987).
31. Sundquist, W. I. & Klug, A. *Nature* **342**, 825–829 (1989).
32. Veselkov, A. G., Malkov, V. A., Frank-Kamenetskii, M. D. & Dobrynin, V. N. *Nature* **364**, 496 (1993).
33. Williamson, J. R. *Rev. Biophys. biomol. Struct.* **23**, 703–730 (1994).
34. Gottschling, D. E. & Zakian, V. A. *Cell* **47**, 195–205 (1986).
35. Price, C. M. & Cech, T. R. *Biochemistry* **28**, 769–774 (1989).
36. Lustig, A. J., Kurtz, S. & Shore, D. *Science* **250**, 549–553 (1990).
37. Larson, G. P., Castanotto, D., Rossi, J. J. & Malafa, M. P. *Gene* **150**, 35–41 (1994).
38. Gottschling, D. E., Aparicio, O. M., Billington, B. L. & Zakian, V. A. *Cell* **63**, 751–762 (1990).
39. Köhrer, K. & Domdey, H. *Meth. Enzym.* **194**, 398–405 (1991).
40. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989).
41. Wray, L. V., Witte, M. M., Dickson, R. C. & Riley, M. I. *Molec. cell. Biol.* **7**, 1111–1121 (1987).
42. Sikorski, R. S. & Boeke, J. D. *Meth. Enzym.* **194**, 302–318 (1991).
43. Church, G. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1991–1995 (1984).

ACKNOWLEDGEMENTS. We thank T. Boswell for help with preliminary experiments, J. Prescott for polyadenylated *K. lactis* RNA, and K. Kirk, A. Krauskopf, M. Mayorga and J. Prescott for comments on the manuscript. This work was supported by a grant from the NIH. Laboratory facilities were provided in part by the Lucille P. Markey Charitable Trust.

LETTERS TO NATURE

Continued decline of total ozone over Halley, Antarctica, since 1985

A. E. Jones & J. D. Shanklin

British Antarctic Survey, Natural Environment Research Council,
High Cross, Madingley Road, Cambridge CB3 0ET, UK

IN 1985, Farman *et al.*¹ announced that a dramatic reduction in total ozone was occurring in the atmosphere over Halley, Antarctica, during the polar spring. Analysis of satellite data revealed that this ozone depletion was an Antarctic-wide phenomenon². Combined theoretical^{3–5}, observational^{6,7} and laboratory⁸ work has shown that chlorine radicals derived from the photolysis of chlorofluorocarbons were the dominant cause of the ozone loss^{9–11}. Ten years later, we review here the status of the 'ozone hole' based on the continued total-ozone measurements at Halley. The springtime 'ozone hole' continues to deepen, with both the October mean and minimum total ozone persistently decreasing. The ozone loss extends into January and February, so that significant increases in ultraviolet-B radiation can be expected at the surface over Antarctica during the summer. A signal of ozone loss is now apparent in the spring and summer temperature records, with recent temperatures at the 100-mbar level consistently close to, or colder than, the historical (1957–72) minima for the period October to January. These low temperatures may well enable the maintenance of springtime ozone-loss mechanisms until later in the year.

Ground-based observations to monitor total ozone have been performed at Halley (75° 31' S, 26° 40' W) since 1957 (ref. 12). The observations of Farman *et al.*¹ showed that the maximum decrease in total ozone occurred in the month of October. They reported a reduction from the historical monthly mean of ~300 Dobson units (DU) to a monthly mean of only 180 DU in 1984. This was well outside the natural variability in the rest of the data record. In Fig. 1a we present the updated version of the Farman *et al.* October mean total ozone plot; Fig. 1b shows the lowest daily total ozone observed over Halley in October each year since 1957 (referred to as the October minimum). It is striking that the sharp decrease in October mean total ozone reported by Farman *et al.* persists throughout the 1980s and into the 1990s. Calculating a least-squares fit to the data between 1976 and 1991 yields a trend of -7.28 ± 2.68 DU yr⁻¹ at the 95% confidence level. In a similar manner, the October minimum total ozone also continues to decrease with no sign of a reduction in the rate of decline; the trend between 1970 and 1980 was

-4.94 ± 1.49 DU yr⁻¹, increasing to -7.75 ± 0.95 DU yr⁻¹ between 1980 and 1990. This continued reduction in the minimum value implies that the declining monthly average is not purely a function of total ozone staying low for longer, but also of additional destruction of the vertical ozone column. It is also interesting to note that the decline in the daily minimum first appeared around 1970, several years before the signal of ozone loss appeared in the monthly mean.

The vertical extent of the 'ozone hole' is not fixed but can vary according to the temperature and stability of the vortex¹³, and in response to the presence of an enhanced sulphate-aerosol loading following volcanic eruptions^{14,15}. The extremely low total ozone observed in October between 1992 and 1994 resulted from the enhanced aerosol loading from the Mount Pinatubo eruption^{29,30} and the particularly cold and stable vortex of 1993 exacerbating the effect of the increasing stratospheric chlorine loading. In October 1993, the monthly mean total ozone was the lowest ever at only 112 DU—a reduction of 63% compared with the historical average. In October of both 1993 and 1994, there were days on which total ozone was less than 100 DU.

In addition to the persistent decline in total ozone observed during the Antarctic spring, there is evidence that reduced levels of total ozone are now extending into late summer. Values of monthly mean total ozone at Halley in January and February are shown in Fig. 2 with a gradual reduction apparent since the late 1970s. These observations support three-dimensional model calculations^{16–18} which have shown that reduced levels of ozone can be expected at high latitudes during the summer as a residue of the springtime 'ozone hole'. As springtime ozone continues to decline there are thus important implications for the amount of ultraviolet-B radiation reaching the Earth's surface during the months with greatest sunlight¹⁹. We calculated trends using least-squares fits to the Halley data. For January, a trend between 1976 and 1991 of -1.05 ± 1.13 DU yr⁻¹ is calculated at the 95% confidence level (significant at the 90% level); between 1976 and 1995, ignoring the years 1992 to 1994 (the Pinatubo years), the trend is -1.05 ± 0.94 DU yr⁻¹ at the 95% confidence level. For February, trends are not yet statistically significant (-0.58 ± 1.41 DU yr⁻¹ at the 95% confidence level between 1976 and 1991). However, 80% of the February monthly means measured in the past 15 years are lower than any observed in the first 15 years of the record. Anomalously low values of ozone were observed in January and February between 1992 and 1994. These coincide with the presence of Pinatubo aerosol over Antarctica, and suggest some *in situ* summertime ozone loss through reactions on the sulphate aerosols in addition to the influence of reduced amounts of springtime ozone. By 1995, following the return of the aerosol loading to roughly pre-

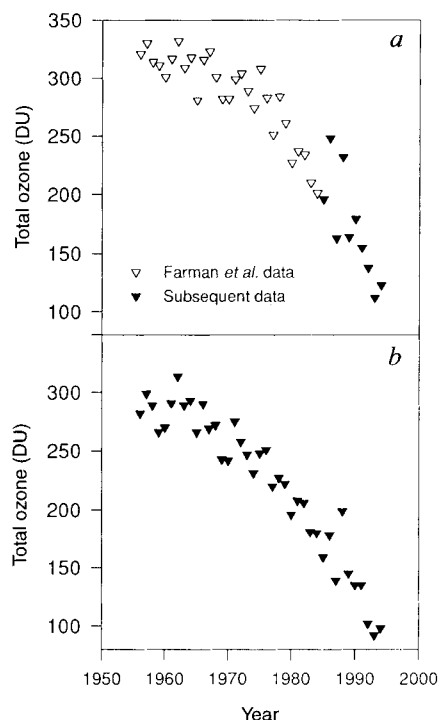


FIG. 1 *a*, October mean total ozone observed over Halley, Antarctica from 1956 to 1994. Empty triangles, data shown in Fig. 2*a* of Farman *et al.*¹; filled triangles, subsequent data. *b*, The lowest daily value of total ozone observed over Halley in October for the years between 1956 and 1994 (the October minimum). The daily value is the mean of all the readings taken during the day (between 5 and 30).

Pinatubo levels, January and February mean total ozone had returned to values expected from the 1980s trend.

Modelling calculations^{16,18,20,21} have suggested that the loss of ozone during the Antarctic spring should cool the lower stratosphere. Temperature soundings have been launched on a daily basis from Halley since 1957. We have considered lower-stratosphere temperature in terms of the variability observed before and after the onset of the 'ozone hole', taking the temperature extremes observed between 1957 and 1972 as the historical bounds. For the period between March and mid-September, all of the variability in lower-stratosphere temperature recorded since the appearance of the 'ozone hole' lies within that previously observed. However, this is not the case for the remaining months. Figure 3 shows the changes in temperature at the 100-mbar level (~15 km height) during the spring and summer. Temperatures were close to the historical mean throughout the 1970s. Some cooling is evident in the early 1980s, and strong departures from the historical range are seen in the late 1980s and the 1990s. The now large decrease in temperature during November is apparent even in the early 1980s. The duration of unusually low temperatures has extended with time, with temperatures below the historical minima recorded frequently in early October and extending into December since the late 1980s. The extremely low temperatures of 1987/88, 1990/91 and 1993/94 coincide with the westerly phase of the equatorial quasi-biennial oscillation (QBO) which characterises deep polar vortices²². Early one-dimensional model radiative calculations²⁰ suggested that the Antarctic ozone loss observed in the mid-1980s should cool the lower stratosphere by more than 5 K in October. Subsequent calculations run in three-dimensional models^{16,18,21}, including both radiative and dynamical considerations, were in good agreement with this early work. Given that any cooling observed in our data is superimposed on the natural interannual tempera-

ture fluctuations, the calculations agree well with temperature changes observed over Halley. In addition, the models showed that the largest temperature decreases would be expected during November, as is observed. Both the magnitude of temperature changes and the timing of the maximum cooling imply that the extremely low temperatures recorded over Halley since the 1980s include a component from ozone loss.

It has been postulated that positive feedbacks between ozone loss and lower-stratosphere temperature might occur within a spring season: the loss of ozone would reduce springtime temperatures, which might prolong the presence of polar stratospheric clouds (PSCs) and consequently that of reactive chlorine, hence extending the period of ozone loss^{23,24}. Even with the now reduced temperatures of the lower stratosphere, the much quoted temperature threshold (~195 K) for the existence of nitric acid trihydrate PSCs could only be observed until late October (note, however, that this threshold depends on HNO_3 and water-vapour mixing ratios). However, there is at present considerable uncertainty as to the exact nature of PSCs. Other solid hydrates of nitric acid, solid hydrates of sulphuric acid, supercooled liquids of nitric acid, and ternary solutions of sulphuric acid, nitric acid and water are all possible components^{25,26}. Consequent differences in PSC threshold temperatures mean that chlorine activation may be possible at warmer temperatures. The uncertainty in PSC composition, however, makes it difficult to assess how late in the year such a feedback would be important. Additionally, a signal of the enhanced greenhouse effect manifesting as a cooling of the stratosphere³¹ could further strengthen the ozone/temperature feedback and further exacerbate ozone loss.

The natural variability of the atmosphere, including local dynamics, the phase of the QBO²⁷, and the occurrence of volcanic eruptions, may influence the severity of Antarctic ozone depletion from year to year. This influence is achieved, however, by exacerbating chlorine-catalysed ozone destruction (apart from any variations in the direct transport of ozone, such as that

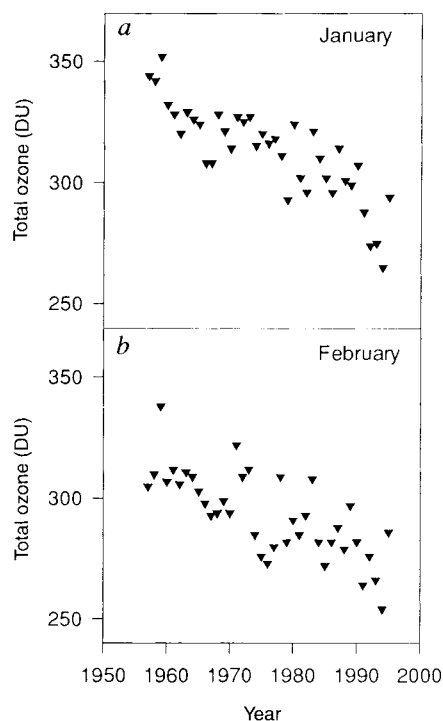
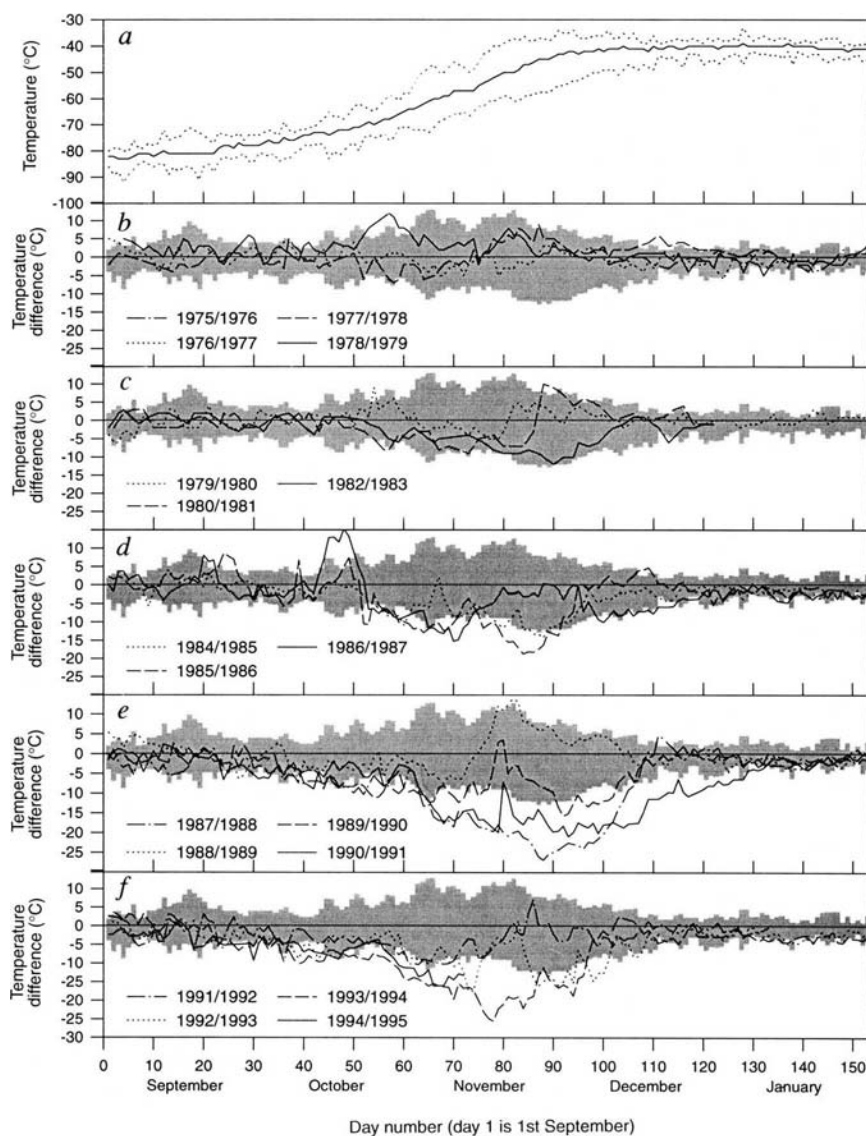


FIG. 2 Monthly mean total ozone observed over Halley from 1957 to 1995 in January (*a*) and February (*b*).

FIG. 3 a, The solid line represents noon temperatures at the 100-mbar level over Halley observed from 1 September to 31 January averaged between 1957 and 1972. Dotted lines mark the maximum and minimum temperature extremes over this time period. b–f, The shaded area represents the 100-mbar temperature extremes shown in a normalized to a historical mean of zero. Superimposed on this are noon 100-mbar temperatures recorded over Halley between 1 September and 31 January for all years between September 1975 and November 1994 (apart from 1981–82 and 1983–84 for which there are no data). Data lying beyond the shaded area lie outside the historical window of extreme temperatures.



associated with the QBO). The seasonal feedback mechanisms discussed above can also amplify the chlorine-catalysed ozone loss. It is currently expected that the chlorine loading of the stratosphere will peak in the late 1990s²⁸. Assuming international compliance with the Copenhagen Amendment²⁸, the chlorine

loading should then fall to the 2 p.p.b.v. threshold (at which the 'ozone hole' first appeared) in the middle of the next century. It will be important to continue to monitor the status of Antarctic ozone during this period to ensure that it does indeed recover to its previous levels. □

Received 20 April; accepted 10 July 1995.

- Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* **315**, 207–210 (1985).
- Stolarski, R. S. et al. *Nature* **322**, 808–811 (1986).
- Solomon, S., Garcia, R. R., Rowland, F. S. & Wuebbles, D. J. *Nature* **321**, 755–758 (1986).
- McElroy, M. B., Salawitch, R. J., Wofsy, S. C. & Logan, J. A. *Nature* **321**, 759–762 (1986).
- Molina, L. T. & Molina, M. J. *J. phys. Chem.* **91**, 433–436 (1986).
- Anderson, J. G., Brune, W. H. & Proffitt, M. H. *J. geophys. Res.* **94**, 11465–11479 (1989).
- de Zafra, R. L. et al. *Nature* **328**, 408–411 (1987).
- Hayman, G. D., Davies, J. M. & Cox, R. A. *Geophys. Res. Lett.* **13**, 1347–1350 (1986).
- Solomon, S. *Nature* **347**, 347–354 (1990).
- Scientific Assessment of Stratospheric Ozone: 1989 (Rep. No. 20, World Meteorological Organization Global Ozone Research and Monitoring Project, WMO, Geneva, 1990).
- Scientific Assessment of Ozone Depletion: 1991 (Rep. No. 25, World Meteorological Organization Global Ozone Research and Monitoring Project, WMO, Geneva, 1992).
- Farman, J. C. & Hamilton, R. A. Scientific Rep. No. 90 (British Antarctic Survey, Cambridge, UK, 1975).
- Hofmann, D. J., Harder, J. W., Rosen, J. M., Hereford, J. V. & Carpenter, J. R. *J. geophys. Res.* **94**, 16527–16536 (1989).
- Deshler, T., Hofmann, D. J. & Hereford, J. V. *J. geophys. Res.* **95**, 10023–10035 (1990).
- Hofmann, D. J. & Oltmans, S. J. *J. geophys. Res.* **98**, 18555–18561 (1993).
- Cariolle, D., Lasserre-Bigory, A., Royer, J.-F. & Geleyn, J.-F. *J. geophys. Res.* **95**, 1883–1898 (1990).
- Prather, M., Garcia, M. M., Suozzo, R. & Rind, D. *J. geophys. Res.* **95**, 3449–3471 (1990).
- Mahlman, J. D., Pinto, J. P. & Umsheid, L. J. *J. atmos. Sci.* **51**, 489–508 (1994).

- Booth, C. R. & Madronich, S. in *Ultraviolet Radiation in Antarctica: Measurements and Biological Research* (eds Weiler, C. S. & Penhale, P. A.) 39–42 (Antarctic Res. Ser. 62, American Geophysical Union, Washington DC, 1994).
- Shine, K. P. *Geophys. Res. Lett.* **13**, 1331–1334 (1986).
- Kiehl, J. T., Boville, B. A. & Briegleb, B. P. *Nature* **332**, 501–504 (1988).
- van Loon, H. & Labitzke, K. *Mon. Weath. Rev.* **115**, 357–369 (1987).
- Poole, L. R., Solomon, S., McCormick, M. P. & Pitts, M. C. *Geophys. Res. Lett.* **16**, 1157–1160 (1989).
- Nichol, S. E., Coulmann, S. & Clarkson, T. S. *Geophys. Res. Lett.* **18**, 1865–1868 (1991).
- Toon, O. B. & Tolbert, M. A. *Nature* **375**, 218–221 (1995).
- Garslaw, K. S. et al. *Geophys. Res. Lett.* **21**, 2479–2482 (1994).
- Garcia, R. R. & Solomon, S. *Geophys. Res. Lett.* **14**, 848–851 (1987).
- Solomon, S. & Wuebbles, D. in *Scientific Assessment of Ozone Depletion: 1994* (Rep. No. 37, WMO, Geneva, 1995).
- Hofmann, D. J. & Oltmans, S. J. *J. geophys. Res.* **98**, 18555–18561 (1993).
- Hofmann, D. J., Oltmans, S. J., Lathrop, J. A. & Vömel, H. *Geophys. Res. Lett.* **21**, 421–424 (1994).
- Fels, S. B., Mahlman, J. D., Schwarzkopf, M. D. & Sinclair, R. W. *J. atmos. Sci.* **37**, 2265–2297 (1980).

ACKNOWLEDGEMENTS. We thank all of the Antarctic personnel who made the measurements presented here; B. G. Gardiner for his work with the Dobson data; S. R. Colwell for providing the temperature data; N. R. P. Harris and E. W. Wolff for useful discussions and comments on the manuscript; S. F. Jones for discussion at the early stages of the work; and the U.K. Meteorological Office for assisting with the radiosonde programme and general use of their facilities.

Preservation of Miocene glacier ice in East Antarctica

David E. Sugden*, David R. Marchant†, Noel Potter, Jr.‡, Roland A. Souchez§, George H. Denton†||, Carl C. Swisher III¶ & Jean-Louis Tison§

* Department of Geography, University of Edinburgh, Edinburgh EH8 9XP, UK

† Institute for Quaternary Studies, || Department of Geological Sciences, University of Maine, Orono, Maine 04469, USA

‡ Department of Geology, Dickinson College, Carlisle, Pennsylvania 17013, USA

§ Département des Sciences de la Terre et de l'Environnement, CP 160/03, Université Libre de Bruxelles, B-1050 Bruxelles, Belgium

¶ Berkeley Geochronology Center, Berkeley, California 94709, USA

ANTARCTIC climate during the Pliocene has been the subject of considerable debate. One view holds that, during part of the Pliocene, East Antarctica was largely free of glacier ice and that vegetation survived on the coastal mountains^{1–4}. An alternative viewpoint argues for the development of a stable polar ice sheet by the middle Miocene, which has persisted since then^{5–10}. Here we report the discovery of buried glacier ice in Beacon valley, East Antarctica, which appears to have survived for at least 8.1 million years. We have dated the ice by ⁴⁰Ar/³⁹Ar analysis of volcanic ash in the thin, overlying glacial till which, we argue, has undergone little (if any) reworking. Isotope and crystal fabric analyses of the ice show that it was derived from an ice sheet. We suggest that stable polar conditions must have persisted in this region for at least 8.1 million years for this ice to have avoided sublimation.

Beacon valley (77° S, 161° E) is near the peripheral Taylor dome of the East Antarctic ice sheet. The present-day mean

annual temperature is –30 to –35 °C and precipitation is less than 10 mm water equivalent per year^{11,12}. The valley is bordered by mountains rising to 3,011 m (Mount Feather), which consist of late Palaeozoic and early Mesozoic sandstones and dolerite (Fig. 1a). The mouth of Beacon valley is dammed by a lobe of Taylor glacier, an outlet glacier that drains Taylor dome on the periphery of the East Antarctic ice sheet. Ablation on Taylor glacier is entirely from sublimation and averages ~0.18 m water equivalent per year at 1,000 m elevation¹³. The glacier is cold-based in its upper ablation zone, although part of the basal ice beneath the lower ablation zone may be at the pressure melting point¹³. Modern $\delta^{18}\text{O}$ values for ice on Taylor dome range^{14,15} from –40‰ to –42‰ (with respect to the SMOW standard); in contrast, $\delta^{18}\text{O}$ values for local glaciers at the head of Beacon valley range from –30‰ to –40‰, reflecting a mix of local snowfall and wind-drifted snow from the ice sheet. (See Fig. 2 legend for definition of $\delta^{18}\text{O}$.) These local glaciers lead down-valley into extensive debris-covered rock glaciers. In between the rock glaciers and Taylor glacier at an altitude of 1,300–1,450 m is a flattish boulder-covered valley floor. It is in this latter area that the buried ice extends over an area of at least 4 km². The thickness of the ice is unknown. But the amplitude of the stagnant ice surface morphology, such as kettle holes, implies a depth of at least 15 m.

The till overlying the buried ice is derived from Taylor glacier and is part of an extensive granite-bearing deposit marking the former expansion of East Antarctic outlet glaciers into the Dry valleys^{5,6,10}. The presence of granite shows the till is not local because this lithology does not crop out in the Beacon valley area and is absent from the deposits of local rock glaciers. Excavation revealed that beneath a surface layer of stones the till is 50 cm thick and consists of a thin upper sandy layer, with angular fragments of local sandstone and dolerite, and a thicker, lower, fine-grained layer containing granite and glacially striated stones in a silt and clay-sized matrix.

TABLE 1 Isotope data and ages of ash deposits from Beacon valley

Analysis no.	⁴⁰ Ar/ ³⁹ Ar	³⁷ Ar/ ³⁹ Ar	³⁶ Ar/ ³⁹ Ar	⁴⁰ Ar*/ ³⁹ Ar	⁴⁰ Ar* (%)	Age (Myr)	s.d.
Sample DMS91-107 (sanidine)							
7182-02	12.5394	1.55483	0.01119	9.368	74.6	8.055	0.293
7182-04	10.7856	0.02569	0.00474	9.387	87.0	8.071	0.222
7182-05	10.4854	0.01234	0.00372	9.387	89.5	8.072	0.184
7182-06	9.5960	0.04598	0.00133	9.206	95.9	7.916	0.158
7182-08	10.4079	0.00000	0.00368	9.319	89.5	8.013	0.477
7182-09	9.6967	0.07168	0.00089	9.439	97.3	8.116	0.080
7182-10	10.7869	0.13244	0.00477	9.389	87.0	8.073	0.208
					Weighted mean	8.07	± 0.06
					Arithmetic mean	8.05	± 0.07
Sample NPS88-2 (sanidine)							
2497-01	4.4866	0.06351	0.00638	2.985	61.3	8.026	0.402
2497-02	3.1188	0.04201	0.00051	2.968	95.2	7.981	0.083
2497-03	7.9987	0.23502	0.01614	3.247	40.6	8.729	0.346
2497-04	3.1251	0.04416	0.00117	2.779	88.9	7.475	0.123
2497-08	3.1431	0.02010	0.00097	2.855	90.8	7.676	0.093
2497-09	3.3614	0.15182	0.00188	2.817	83.8	7.577	0.170
2497-10	3.1806	0.03697	0.00120	2.826	88.9	7.601	0.196
					Weighted mean	7.87	± 0.43
					Arithmetic mean	7.78	± 0.05
Contaminant or detrital grain							
2497-07	4.7721	0.04123	0.00153	4.323	90.6	11.611	0.191

Laser total-fusion ⁴⁰Ar/³⁹Ar analyses of individual euhedral crystals from ash deposits overlying buried ice in Beacon valley, carried out by the Berkeley Geochronology Center, Berkeley, California. Pristine euhedral crystals were hand-picked from each ash sample using a binocular microscope, treated with 7% hydrofluoric acid in an ultrasonic cleaner for 5 min to remove any altered clays or attached glass, followed by 10 min in distilled water, and then irradiated in the core of the Omega West research reactor at Los Alamos National Laboratory. For sample DMS-91-107, obtained in 1991 from the wedge illustrated in Fig. 1b, the weighted mean age of 8.07 ± 0.06 Myr is based on seven separate single-crystal analyses using a Mass Analyzer Product 215 noble-gas mass spectrometer, calibrated with monitor minerals Fish Canyon Sanidine and MMhb-1; the weighted mean age of 7.87 ± 0.43 Myr for sample NPS88-2 is also based on seven separate analyses. This latter sample was collected from an adjacent relic wedge in 1988. Asterisk describes radiogenic ⁴⁰Ar.

The till surface is marked by contraction-crack polygons¹⁶. Active forms are characterized by V-shaped troughs 2–3 m deep, which penetrate into the underlying ice and are commonly filled with wedges of vertically stratified sand and gravel deposits. In addition there are traces of relict wedges, recognizable only in stratigraphic section in the overlying till, which reflect abandoned crack patterns. We have discovered volcanic ash in four such relict wedges. In one well preserved wedge shown in Fig. 1b, the ash has a uniform geochemistry and contains <10% non-volcanic contaminants. Its grain size distribution is bimodal with peaks at 23 and 250 μm but with a coarse component up to 1,000 μm in size. The ash fines upwards within the wedge. It contains glass shards with fragile bubble vesicles and delicate spires. There is no chemical etching of volcanic anorthoclase

crystals, and glass shards show no evidence of weathering when viewed under a scanning electron microscope.

The high concentration of ash, the uniform geochemistry, the bimodal grain-size distribution, the stratification by grain size, and the preservation of delicate features on glass shards all point to an origin by direct airfall into an open crack¹⁷. Moreover, the preservation of details of the wedge structure and the lack of mixing and abrasion of ash particles suggests that there has been little subsequent disturbance. These observations provide compelling grounds for arguing that the ash is in situ, apart from any settling associated with sublimation of the underlying ice. Further, on the basis of examination of 75 volcanic ash sites in the area, we can preclude alternative explanations for the stratigraphic position of the ash¹⁷. Aeolian action produces an

abraded, unimodal deposit mixed with non-volcanic material, whereas down-slope mass movement or rock glacier flow distorts the initial wedge form, and likewise causes grain abrasion and mixing.

Laser-fusion $^{40}\text{Ar}/^{39}\text{Ar}$ dating of individual crystals shows that the ashes are 7.9–8.1 Myr old (Table 1). This means that the granite-bearing till in which they are emplaced must be older than 8.1 Myr, a conclusion which agrees with other minimum dates for the same till elsewhere in the Dry valleys, also obtained by isotopic dating of volcanic ashes. The latter yield minimum ages of 11.3 Myr in the adjacent Arena valley, and 13.6 and 12.0 Myr in two separate locations in the Asgard Range^{5,6}. The ash dates imply that the granite-bearing till is Miocene or older in age.

The surface of the ice beneath the ash wedge is approximately horizontal, and truncates a plunging fold (Fig. 1c). The ice, which consists of equigranular 10-mm crystals, is clouded by dispersed rock debris, and contains several debris bands up to 150 mm in thickness (Fig. 2a). Granulometric analyses indicate that the debris is composed of 3% gravel (a small percentage of which is granite), 72% sand and 15% silt and clay. Enveloping the debris-bearing layers are bands of clear ice and white, bubbly foliated ice. Alternating debris-rich and debris-poor sequences are typical of basal regelation and debris entrainment, and reflect the former presence of basal ice at the pressure melting point^{18,19}. Detailed laboratory analysis of thin sections from each of the ice types reveals a fabric with a strong maximum caused by shearing at depth with some compressive flow (Fig. 2b). The preservation of this type of fabric and the small crystal size may reflect the ability of dirt content to impede recrystallization processes^{20,21}. The direction of movement is up-valley and oblique to the longitudinal axis of Beacon valley; this is consistent with the presence of granite, and shows that the ice is derived from an expanded Taylor outlet glacier rather than local glaciers.

Oxygen-isotope analyses of the fossil ice give $\delta^{18}\text{O}$ values that cluster around -33‰ . They yield less negative values than exist at higher altitudes on Taylor dome today. Because $\delta^{18}\text{O}$ values for the fossil ice fall within the range obtained for modern ice at the head of Beacon valley, we suggest that

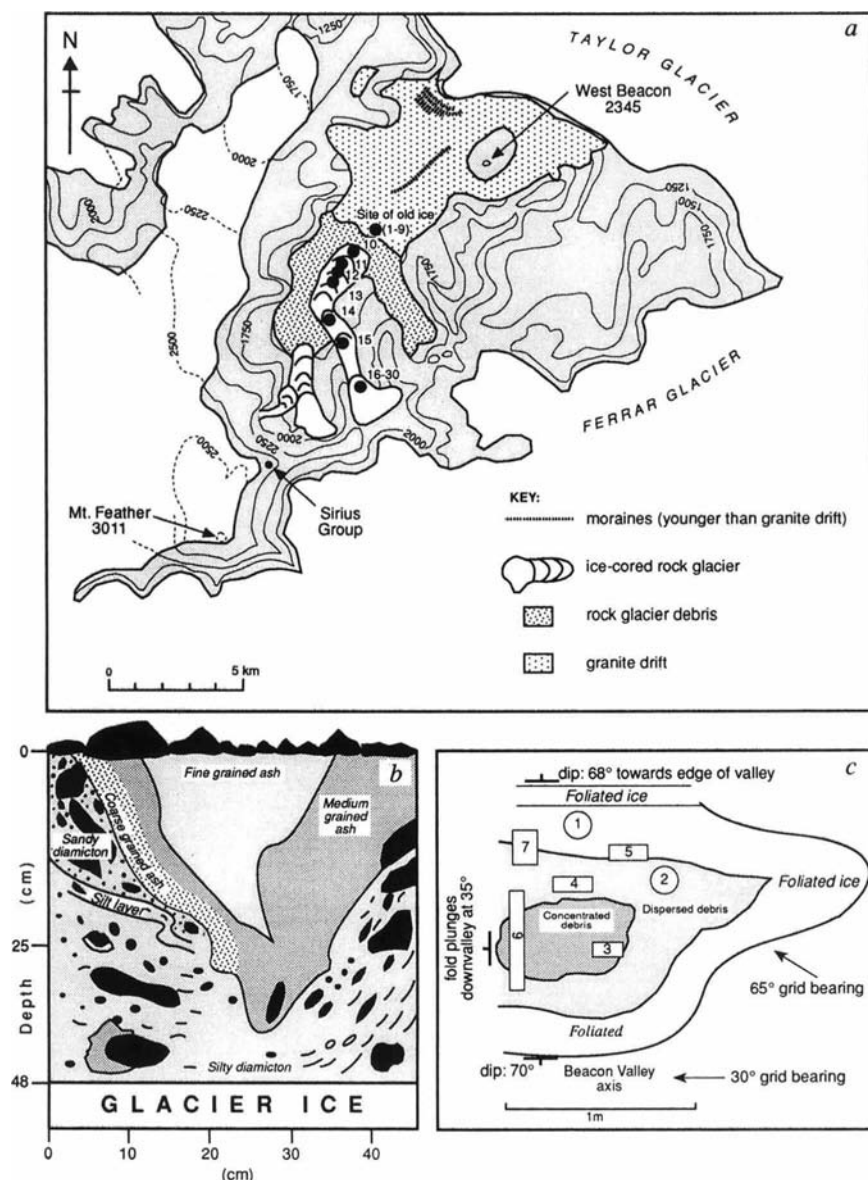


FIG. 1 a, Map showing Beacon valley and the sites (filled circles, numbered 1–30) where ice was sampled. Heights are shown in metres. The old ice underlies much of the granite drift on the floor of Beacon Valley. b, Excavation showing the vertical relationships between the till, the wedge of volcanic ash and the underlying ice. The ash-filled wedge, which tapers with depth, is bounded by vertically oriented stones and has been little disturbed since formation. The lower, granite-bearing, silty diamicton has a matrix with 1.7 wt% clay-sized particles (<0.004 mm). The coarse stratification of the ash by size is typical of a direct airfall deposit. c, Plan view of the ice exposed in the excavation beneath the ash wedge, showing sampling sites and how the surface truncates a plunging fold. Ice was transferred, without melting, to Brussels for analysis.

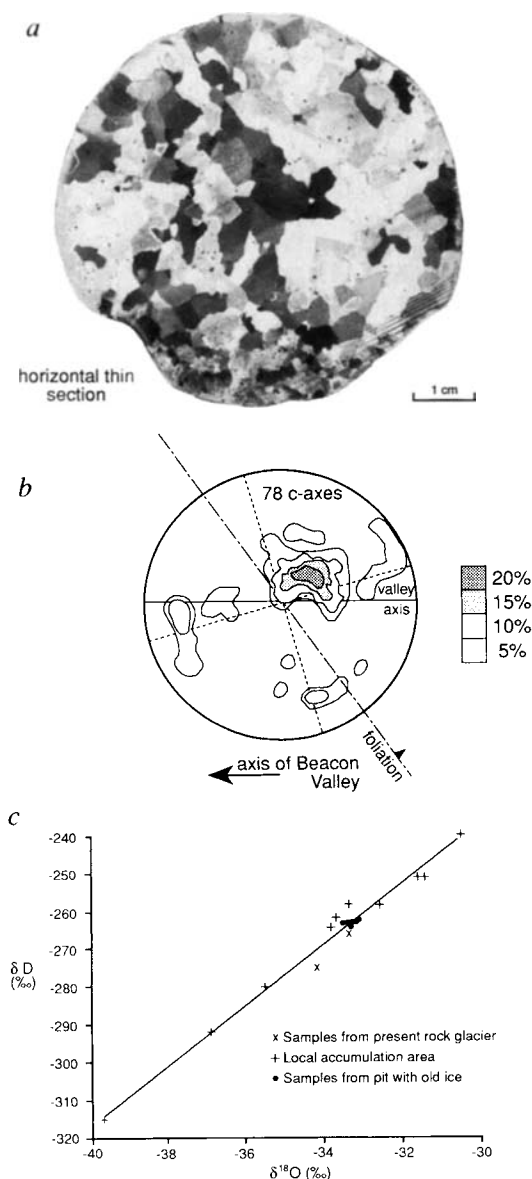


FIG. 2 Details of the analysis of the fossil ice and entrained debris. *a*, Thin section showing the ice crystals (~ 10 mm in size) and bubbles in a sample from site 2 (Fig. 1c). *b*, Stereographic projection showing the orientations of 78 c-axes in ice in the outer limb of the fold. This reveals a strong fabric reflecting ice moving obliquely up-valley. Density of c-axes is expressed as the number of points per one per cent area. *c*, Stable-isotope analyses, showing the contrast between the cluster of values derived from the old ice in the pit with the spread of values from the active rock glacier and from the accumulation area at the head of Beacon valley. All values fall along the precipitation slope and there is no sign of refreezing. ($\delta^{18}O = [(^{18}O/^{16}O)_{\text{sample}} / (^{18}O/^{16}O)_{\text{SMOW}}] - 1$ per mil and $\delta D = [(D/H)_{\text{sample}} / (D/H)_{\text{SMOW}}] - 1$ per mil.)

climate conditions on the ice dome of the Miocene-age ice were similar to those of Beacon valley today. All ice samples, fossil and modern, are aligned on a precipitation slope, implying that phase changes in the fossil ice have not significantly modified the initial isotopic composition.

Overall, the debris and ice characteristics suggest that this is a remnant of basal glacier ice. The nature of the debris shows that the glacier was sliding over its bed; the ice fabric and the presence of granite erratics suggest that the remnant glacier ice was derived from an expanded Taylor glacier. The isotopic analyses suggest that the ice accumulated under cold climate

conditions similar to those of Beacon valley today. These findings all point to the preservation of a remnant of an ancient outlet glacier of the East Antarctic ice sheet for at least 8.1 Myr.

At first sight it might seem surprising that glacier ice can avoid sublimation beneath a thin layer of till over a period of 8 Myr. However, if one assumes that the soil is 100% saturated with water vapour and that the water partial-pressure gradient is determined on average by the geothermal heat flux²², it is found that the amount of sublimation is consistent with 1 m of settling. A problem arises in that relative humidities below 100% have been observed in soils elsewhere in the Dry valleys²³; if these lower values applied to Beacon valley then they could lead to moisture fluxes three orders of magnitude greater. At this stage, when little is understood about the role of sublimation in frozen sediments²⁴, it is useful to highlight three points in support of our assumptions: the observed relative humidity fluctuations²³ were restricted to near the surface, and saturation persisted at a depth of 30–40 cm; the readings are atypical in that they were taken in summer when surface temperatures were highest and relative humidities lowest; the presence of snow causes the underlying soil to be saturated, and in Beacon valley drifting snow collects preferentially in the depressions associated with wedges. We conclude that lowering of a few metres is plausible. It is consistent with the preservation of the overlying ash-wedge structure and would also explain the loss by settlement of any traces of the lower part of the wedge which may originally have been 2–3 m deep.

The implication of this discovery of ancient glacier ice is that a polar ice sheet, sufficiently big to nourish outlet glaciers flowing into the Dry valleys, existed in East Antarctica by 8.1 Myr ago. There is a further implication for Pliocene climate. The survival of ancient glacier ice beneath a thin layer of unconsolidated till suggests that the cold polar climate of Beacon valley has persisted at least for the past 8.1 Myr, and supports those who argue for the long-term stability of the East Antarctic ice sheet^{5, 10}.

Received 2 September 1994; accepted 10 July 1995.

- Webb, P. N. & Harwood, D. M. *Antarct. J. U.S.* **22**, 7–11 (1987).
- Webb, P. N., Harwood, D. M., McKelvey, B. C., Mercer, J. H. & Stott, L. D. *Geology* **12**, 287–291 (1984).
- Barrett, P. J., Adams, C. J., McIntosh, W. C., Swisher, C. C. & Wilson, G. S. *Nature* **359**, 816–818 (1992).
- Webb, P. N. & Harwood, D. M. in *Forest Development in Cold Climates* (eds Alden et al.) 135–165 (Plenum, New York, 1993).
- Marchant, D. R., Denton, G. H. & Swisher, C. C. *Geogr. Annlr. Stockh.* **75A**, 269–303 (1993).
- Marchant, D. R., Denton, G. H., Sugden, D. E. & Swisher, C. C. *Geogr. Annlr. Stockh.* **75A**, 303–330 (1993).
- Kennett, J. P. *Geogr. Annlr. Stockh.* **75A**, 205–220 (1993).
- Marchant, D. R., Swisher, C. C., Lux, D. R., West, D. P. Jr & Denton, G. H. *Science* **260**, 667–670 (1993).
- Sugden, D. E. *Nature* **359**, 775–776 (1992).
- Denton, G. H., Sugden, D. E., Marchant, D. R., Hall, B. L. & Wilch, T. I. *Geogr. Annlr. Stockh.* **75A**, 155–204 (1993).
- Schwerdtfeger, W. *Developments in Atmospheric Science* Vol. 15 261 pp (Elsevier, Amsterdam, 1984).
- Thompson, D. C., Craig, R. M. F. & Bromley, A. M. *New Zealand J. Sci.* **14**, 245–251.
- Robinson, P. H. *J. Glaciol.* **30**, 153–160 (1984).
- Hendy, C. H., Healy, T. R., Rayner, E. M., Shaw, J. & Wilson, A. T. *Quat. Res.* **11**, 172–184 (1979).
- Drewry, D. J. in *Antarctic Geoscience* (ed. Craddock, C.) 977–983 (Univ. Wisconsin Press, Madison, 1982).
- Black, R. F. *Quat. Res.* **6**, 3–26 (1976).
- Marchant, D. R., Denton, G. H., Swisher, C. C. & Potter, N. Jr *Bull. geol. Soc. Am.* (in the press).
- Souchez, R. A., Lorrain, R., Tison, J.-L. & Jouzel, J. *Ann. Glaciol.* **10**, 163–166 (1988).
- Sugden, D. E. et al. *Nature* **328**, 238–241 (1987).
- Alley, R. B., Perepezko, J. H. & Bentley, C. R. *J. Glaciol.* **32**, 425–493 (1986).
- Tison, J.-L., Thorsteinsson, T., Lorrain, R. D. & Kipfstuhl, J. *Earth planet. Sci. Lett.* **125**, 421–437 (1994).
- Smith, M. W. & Burn, C. R. *Cold Regions Science and Technology* **13**, 143–152 (1987).
- Cameron, R. E. & Conrow, H. P. *Antarct. U.S.* **4**, 23–28 (1969).
- Law, J. & van Dijk, D. *Permafrost and Periglacial Processes* **5**, 237–249 (1994).

ACKNOWLEDGEMENTS. We thank J. Jouzel for the isotope analyses. This work was supported by the US National Science Foundation, the Natural Environment Research Council, the Royal Society of London, the Belgian Antarctic Research Programme, and the Royal Society of Edinburgh.

Size and structure of the Chicxulub crater revealed by horizontal gravity gradients and cenotes

A. R. Hildebrand*, M. Pilkington*, M. Connors†, C. Ortiz-Aleman‡ & R. E. Chavez‡

* Geological Survey of Canada, 1 Observatory Crescent, Ottawa, Ontario K1A 0Y3, Canada

† Department of Physics, University of Alberta, Edmonton, Alberta T6G 2E1, Canada

‡ Instituto de Geofísica, Ciudad Universitaria, Delgacion de Coyoacan, Codigo 04510, Mexico DF, Mexico.

It is now widely believed that a large impact occurred on the Earth at the end of the Cretaceous period, and that the buried Chicxulub structure in Yucatán, Mexico, is the resulting crater²⁴. Knowledge of the size and internal structure of the Chicxulub crater is necessary for quantifying the effects of the impact on the Cretaceous environment. Although much information bearing on the crater's structure is available, diameter estimates range from 170 to 300 km (refs 1–7), corresponding to an order of magnitude variation in impact energy. Here we show the diameter of the crater to be ~180 km by examining the horizontal gradient of the Bouguer

gravity anomaly over the structure. This size is confirmed by the distribution of karst features in the Yucatán region (mainly water-filled sinkholes, known as cenotes). The coincidence of cenotes and peripheral gravity-gradient maxima suggests that cenote formation is closely related to the presence of slump faults near the crater rim.

The size and structure of the buried Chicxulub crater have been primarily constrained from associated circular gravity and magnetic anomalies, although the distribution of karst topography, subsurface drill information and seismic reflection data have also been used^{1–7}. To provide better resolution of Chicxulub's structure we acquired closely spaced (0.5–3 km) gravity measurements, primarily along radial profiles across the crater's rim. Here we present results from integrating 310 stations from the crater's southwestern quadrant with the proprietary data set acquired by Petróleos Mexicanos. The data were gridded to a 750-m interval, and the horizontal gradient⁸ of the Bouguer anomaly was computed. Calculating gravity gradients emphasizes the effect of lateral density changes, and suppresses regional gradients that may obscure the gravity signature of the impact lithologies. This technique provides an objective method of determining where structure lies in the gravity data.

Figure 1 reveals a striking circular structure centred near the Yucatán shoreline at ~89.57° N, 21.29° W. At least six concentric gradient features occur between ~20 and ~90 km radius. These features are probably most distinct in the southwest owing to denser sampling of the gravity field. Truncations of the gradi-

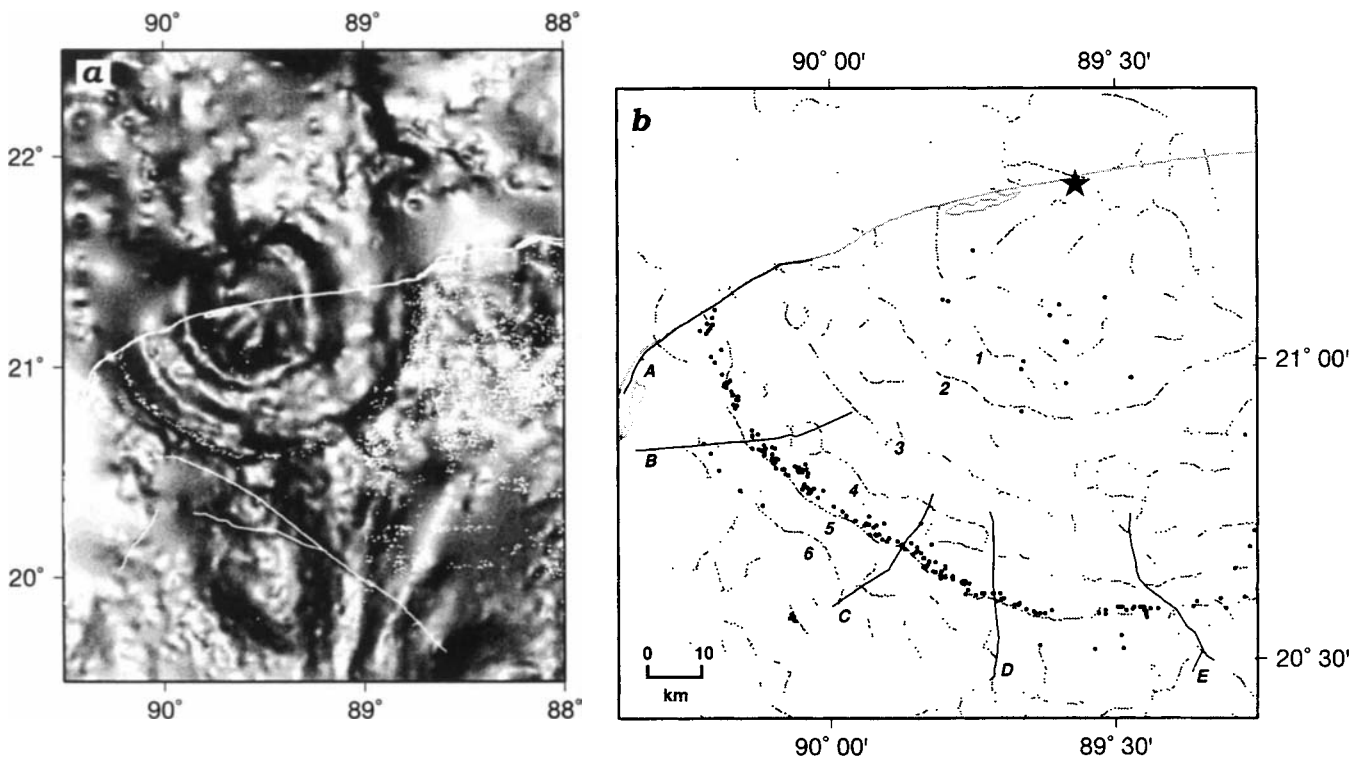


FIG. 1 a, Horizontal gradient of the Bouguer gravity anomaly (unfiltered). The calculation method root-sum-squares the gradient in two orthogonal directions in a 5×5 grid cell (3.75×3.75 km)⁸. Darker shading corresponds to greater gradient magnitude. The technique's sensitivity to subsurface structure is illustrated by the detection of part of the Ticul fault that has only ~100 m of vertical displacement at the surface. The locations of the Yucatán coastline, cenotes and the splays of the Ticul fault are shown in white. Linear east-west and south-north features in the distribution of cenotes are mapping artefacts. Areas lacking gravity data appear blurred, and offshore ship tracks result in discrete linear features. Although the ship tracks are relatively widely spaced so that the gravity field is not as well sampled for the submarine portion of the crater, sea-surface data²³ confirm that the peripheral strong gradient

features are truncated or diverted for the northern third of the crater³. Magnetic² and seismic data^{1,7} confirm that a completely circular basin and impact structure is present, and some weak circular structure appears in the gravity data north of the truncation of the peripheral gradient features, but the cause of the truncation of Chicxulub's gravity expression remains to be understood. b, Detail of the southwestern part of crater showing location of six gradient feature maxima (>0.4 mGal km⁻¹), shown as dots, and karst features (filled circles). The solid lines show segments of five gravity survey profiles (labelled A to E) between 60 and 90 km radius; the coast is shown as a grey line. The star indicates the approximate centre of the crater³. The cenotes occur interior to the gradient maximum, except near the coast where the gridding interpolation is not constrained by any nearby data.

ent features often correspond to gaps in survey coverage. Our detailed profiles detected an additional feature and steeper gradients (up to 5 mGal km^{-1}) than the original survey. We interpret the outer four gradient maxima (at ~ 55 to $\sim 90 \text{ km}$ radii) to represent concentric faults in the crater's zone of slumping, as are also revealed by seismic reflection data⁷. The inner two maxima probably represent the outer margins of the central uplift (at $20\text{--}25 \text{ km}$ radius) and the peak ring and/or collapsed transient cavity (at $40\text{--}45 \text{ km}$ radius). Radial gradients in the southwestern quadrant over the inferred $\sim 40\text{-km}$ -diameter central uplift may represent structural 'puckering' as revealed at eroded terrestrial craters⁹. Gradient features related to regional gravity highs and lows are visible outside the crater^{10,11}, but no concentric gradient features are apparent at radial distances $>90 \text{ km}$. Note that the crater's outer gradient and karst features are linear near the northern and tangential part of the Ticul fault (between profiles B and D, Fig. 1b). This linearity in the zone of slumping, and the more recent vertical displacement along the Ticul fault, probably reflect the presence of northwest-trending structure in the underlying crystalline basement. Slumping at other complex craters has often been controlled by pre-existing structures in the impacted rocks¹².

Figure 1b plots the gravity-gradient maxima and cenote positions for the western part of the crater. Locations of the five radial gravity profiles plotted in Fig. 2 are also shown. The

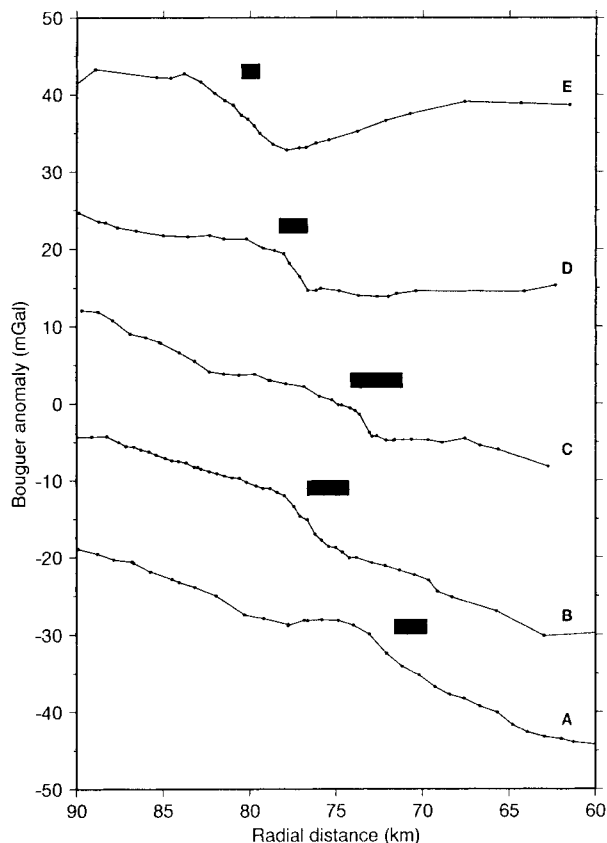


FIG. 2 Five radial profiles from 60 to 90 km radius arranged top to bottom in clockwise order as shown in Fig. 1b. For display purposes, the upper two profiles are offset by adding 25 and 37.5 mGal, respectively; the bottom two are offset by subtracting 25 and 38.75 mGal, respectively. Note that the distance to the steepest gradient feature decreases from southeast to northwest, with the exception of the central profile (C) where the linear portion of the zone of slumping occurs. The overall profile trends vary owing to differing regional fields. The horizontal black bars show the radial extent of cenotes' positions within $\sim 5 \text{ km}$ of the respective profiles.

differences between the profiles are in part due to varying regional fields, but probably also reflect meandering structure in the outer part of the zone of slumping, as typically observed at large complex craters¹². The concentric gradient features probably correspond to density contrasts in the underlying crater rocks, and need not bear any direct relation to the crater's pre-burial topography. The crater's gravity signature can be modelled using only density contrasts between the country rocks and impact lithologies².

Large horizontal gravity gradients require shallow lateral density contrasts, large density contrasts, and/or an extensive near-vertical boundary between units of different density. Available seismic-reflection data include five sections across the crater's slump zone^{1,7} that image displacements on normal faults of up to 0.9 s of two-way travel time (corresponding to $\sim 1.7 \text{ km}$ of vertical offset¹³). The nearest seismic profile lies $\sim 40 \text{ km}$ north of the coastal traverse and shows a normal fault of $\sim 1.5 \text{ km}$ displacement that projects to the location of the outer large gradient feature (no. 5), so we have constructed a model assuming that this gradient maximum reflects a slump fault. Figure 3 shows that a model based on slump faults can reproduce the observed gravity profile including the steep gradient feature and the two adjacent weaker gradient features (nos 4 and 6). The crater rim is interpreted to lie near the outermost feature. The model's outermost slump terrace width is $\sim 9 \text{ km}$, comparable to that predicted by the terrace-width/diameter relationship for lunar craters¹⁴, suggesting that the cratering slump process scales independently of gravity. The 'wiggles' ($10\text{--}15 \text{ km}$ long) in the outer steep gradient maximum (Fig. 1b) probably represent the scalloping at the headwall of the inferred slump terrace, as typically observed at large craters¹². Other structural models for the crater^{5,6} are difficult to reconcile with the observed gravity (and seismic-reflection) data, as they show the zone of slumping beginning at, and extending outwards from, $\sim 85 \text{ km}$ radius where no gradient features (which should be detectable based on their proposed crater geology) are found.

Figure 1 reveals a spatial correlation between the prominent cenote ring associated with Chicxulub⁵ and the outer steep gravity-gradient feature (no. 5). The locations of the karst features plotted in Fig. 1 were digitized from the $1:50,000$ topographic maps of the region, which show most but not all the water-filled cenotes; at the western coastal terminus of the cenote ring, the cenotes' positions were mapped directly from $1:75,000$ scale aerial photographs. The mapped cenotes constitute an unbiased sampling of the region's karst surface features of $>50 \text{ m}$ diameter. Figure 1b shows that the gradient feature (no. 5) generally bounds the outer edge of the cenote ring; this was confirmed by the locations of the steepest gradient found by each of the surveyed profiles. The gradient maximum and the cenote ring both meander with amplitudes of up to 2 km . A second partial cenote ring exterior to the southwestern side of the main ring⁵ corresponds to a less-prominent gravity-gradient feature (no. 6). No concentric structure is observable in the distribution of karst features at radii $>90 \text{ km}$.

The cenote ring corresponds to (and was presumably created by) a zone of enhanced groundwater flow, as evidenced by a coincident low in the groundwater surface¹⁵ and the presence of freshwater springs where the cenote ring intersects the coast⁵. High hydraulic conductivity is also indicated in the surrounding near-surface rocks by low hydraulic gradients¹⁵ and widespread response to tides¹⁶. The zone of dissolution is developed at the top of a sequence of carbonate sediments (of Palaeocene to Pliocene age) of $>0.5 \text{ km}$ thickness at the crater rim¹⁷; the crater units therefore do not directly influence the near-surface groundwater flow. The ring consists mostly of water- and reed-filled cenotes of larger size (typically up to $\sim 300 \text{ m}$ diameter) than the cenotes and dry karst pits found elsewhere on the peninsula. On the crater's east side, the size of the main ring's constituent cenotes allows it to be distinguished from the widespread karst features exterior to the crater. The main cenote ring is up to

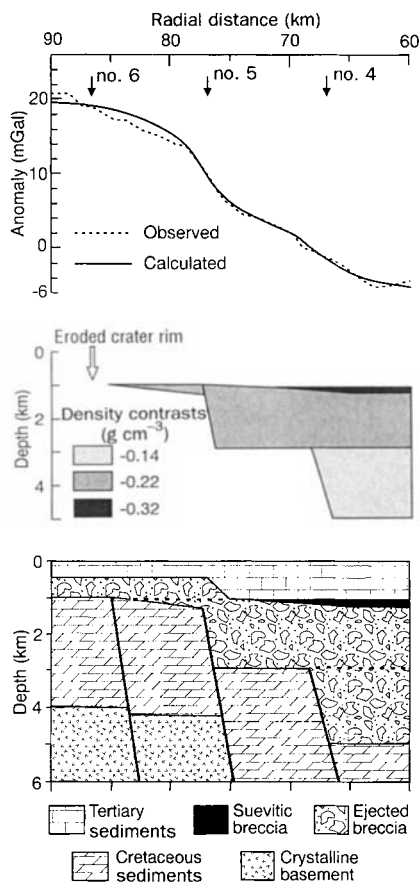


FIG. 3 Density contrast model (middle) and inferred geology (bottom) across the rim of the Chicxulub crater (vertical exaggeration of 3 times). The observed gravity data (profile B; Fig. 1b) were collected along Route 25 between Kinchil and Celestun and are projected onto a crater radius (top). The remainder of the crater gravity model is similar to that of Pilkington *et al.*², except that almost none of the mass deficiency associated with the crater results from lower densities having been induced in the slumped blocks. Note that the density contrasts are between the crater units and rocks located at the same depth outside the crater. For example, the lower half of the crater fill is assumed to have the same density as the proximal ejecta blanket. Dashed lines locate the boundaries of density-contrast polygons where they do not correspond to faults or unit boundaries (bottom). Based on the few data available², we estimate that the densities of the Tertiary sediments are mostly in the range 2.4–2.5 g cm⁻³ which yields plausible densities for the other geological units given the density contrasts used here. The steepness of the peripheral gradient features (indicated by numbered solid arrows) requires the mass deficiencies responsible to be located at relatively shallow depths. Although this model is simplified, it illustrates that the peripheral steep gradient features can be produced by schematic crater structure consistent with the slumping revealed by seismic-reflection studies⁷. Note that we assume that the upper ~0.5 km of Cretaceous sediments were eroded and transported by emplacement of the ejecta blanket.

~5 km wide, and corresponds to a topographic low of up to ~5 m over some of its length. Much of the permeability in karst terrains like the Yucatán peninsula is due to fracture systems¹⁵, so it has been assumed that the cenote ring corresponds to a zone of enhanced fracturing^{5,18}. Although fractures are not directly observable along the cenote ring¹⁵, we presume that a fracture system of ≤5 km width with orientations preferentially parallel to the underlying crater structure created the permeability that lead to the ring's formation. A zone of enhanced fracturing would also allow preferential erosion to create the topographic low associated with the ring.

The correlation of the largest peripheral gradient maximum with the outer bound of the cenote ring (and a weaker feature with the partial cenote ring) strongly indicates a genetic link between the crater's peripheral faults and the assumed fracturing. Previously proposed formation mechanisms for the cenote ring include reactivation of crater slump faults by various means^{5,19,20}. Reactivation of the peripheral slump faults is appealing, but this mechanism needs to account for the 5-km width of the main zone of high permeability and its occurrence interior to the surface projection of the inferred fault. Long-lasting subsidence of at least 200 Myr duration is observed at the ~15-km-diameter Ames crater²¹, so Chicxulub's greater size and 65-Myr age allow for analogous continuing subsidence. Although no displacements have yet been observed in the rocks channelling the ground water¹⁸, the uppermost Tertiary strata are thicker inside the crater than in exterior wells¹⁷, supporting the possibility of fractures having been induced along peripheral faults by continued subsidence. The postulated subsidence could reflect differential compaction or thermal relaxation of the impact lithologies. In addition to the gradient/cenote connection, the buried crater also influences erosion of surface rocks in other ways. The edges of the crater (and associated gravity-gradient features) correspond to bends in the coastline²², and interior gradient features sometimes correlate with rocky points along the coastline.

Gravity data are able to image Chicxulub's structure to an extent unprecedented for this technique (as applied to other terrestrial impact structures), and will guide further exploration of the crater. Additionally, mapping gravity gradients will probably yield considerable insight into groundwater flow on the north-western corner of the Yucatán peninsula, potentially providing considerable societal benefit. The mechanism whereby the buried crater controls near-surface groundwater flow is not yet established, but the gravity-gradient/cenote correlation must figure in the explanation. □

Received 29 March; accepted 10 July 1995.

- Hildebrand, A. R. *et al.* *Geology* **19**, 867–871 (1991).
- Pilkington, M., Hildebrand, A. R. & Ortiz-Aleman, C. *J. geophys. Res.* **99**, 13147–13162 (1994).
- Hildebrand, A. R., Connors, M., Pilkington, M., Ortiz-Aleman, C. & Chavez, R. E. *Rev. Soc. Mex. Paleol.* **7**, 59–68 (1994).
- Espindola, J. M., Mena, M., de La Fuente, M. & Campos-Enriquez, J. O. *Phys. Earth planet. Inter.* (in the press).
- Pope, K. O., Ocampo, A. C. & Duller, C. E. *Earth Moon Planets* **63**, 93–104 (1993).
- Sharpton, V. L. *et al.* *Science* **261**, 1564–1567 (1993).
- Camargo-Zanoguera, A. & Suárez-Reynoso, G. *Bol. Asoc. Mex. Geof. Expl.* **34**, 1–28 (1994).
- Sharpton, V. L., Grieve, R. A. F., Thomas, M. D. & Halpenny, J. F. *Geophys. Res. Lett.* **14**, 808–811 (1987).
- Bunting, J. A., de Laeter, J. R. & Libby, W. G. *Annual Report 1979* 81–85 (Geol. Surv. of Western Australia, Perth, Western Australia, 1980).
- Ness, G. E., Dauphin, J. P., Garcia-Abdeslem, J. & Alvarado-Omana, M. E. *Map and Chart Series MCH073* (Geol. Soc. Am., Boulder, CO, 1991).
- Pilkington, M., Grieve, R. A. F., Rupert, J. D. & Halpenny, J. *Map 1807A* (Geol. Surv. Can., Ottawa, 1992).
- Melosh, H. J. *Impact Cratering: A Geologic Process* 245 (Oxford Univ. Press, New York, 1989).
- Cué, A. V. *Bol. Asoc. Mex. Geol. Petrol.* **5**, 285–290 (1953).
- Pearce, S. J. & Melosh, H. J. *Geophys. Res. Lett.* **13**, 1419–1422 (1986).
- Marin, L. E. thesis, Northern Illinois Univ. (1990).
- De Landa, D. *Yucatán Before and after The Conquest* 184 (trans. by Gates, W., San Fernando, Mérida, México, 1556).
- Lopez-Ramos, E. *The Gulf of Mexico and the Caribbean* (eds Nairn, A. E. M. & Stehli, F. G.) 257–282 (Plenum, New York, 1975).
- Cartas Geológicas, Calkini F15-9-12, Tizimin, F16-7, Merida, F16-10* (Instituto Nacional de Estadística, Geografía e Informática, Mexico City, 1984).
- Pope, K. O., Ocampo, A. C. & Duller, C. E. *Nature* **351**, 105 (1991).
- Perry, E. C., Marin, L., McClain, J. & Velázquez, G. *Geology* **23**, 17–20 (1995).
- Johnson, K. S. & Smith, D. (abstr.) in *Ames Structure and Similar Features—A Workshop* 36 (Oklahoma Geol. Surv., Norman, OK, 1995).
- Hildebrand, A. R. thesis, Univ. Arizona, (1992).
- Sandwell, D. T., Yale, M. M. & Smith, W. H. F. (abstr.) *Eos* **75**, 155 (1994).
- Smit, J. *Nature* **365**, 809–810 (1994).

ACKNOWLEDGEMENTS. We thank D. M. Jobin, P. Salib and C. Chamberlain for helping to reduce the gravity data, the Defense Mapping Agency for supplying the Petróleos Mexicanos gravity data set, W. Roest for supplying the gradient maximum locator program, R. Whistance-Smith who assisted with mapping cenotes, and W. Tremblay for digitizing the coastline. The survey work was supported by the Pan American Institute of Geography and History of the OAS, and PADEP of the Universidad Autónoma de México.

A trade-off-invariant life-history rule for optimal offspring size

Eric L. Charnov* & Jerry F. Downhower†

* Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA

† Department of Zoology, Ohio State University, Columbus, Ohio 43210, USA

OPTIMIZATION models have been widely and successfully used in evolutionary ecology to predict the attributes of organisms¹⁻⁶. Most such models maximize darwinian fitness (or a component of fitness) in the face of trade-offs and constraints; the numerical results usually depend on the exact form of the trade-offs/constraints. Here we report the first (to our knowledge) numerical optimum for life-history evolution which is independent of the details of the underlying trade-off, for a large array for trade-off forms. The rule is that at small litter sizes, the range in offspring size is inversely proportional to the size of the litter. Details of the offspring-survival/offspring-size trade-off⁷⁻¹⁰ set the value of the proportionality constant, but the -1 exponent, the inverse proportionality itself, is universal. Studies of life histories have yielded many empirical examples of universality for various scaling exponents⁶ (for example, adult lifespan scales as ≈ 0.25 with adult body mass within many taxa); this is an example of the numerical value of an exponent (here -1) emerging from a life-history model as independent of all but a few general features of the underlying economic structure.

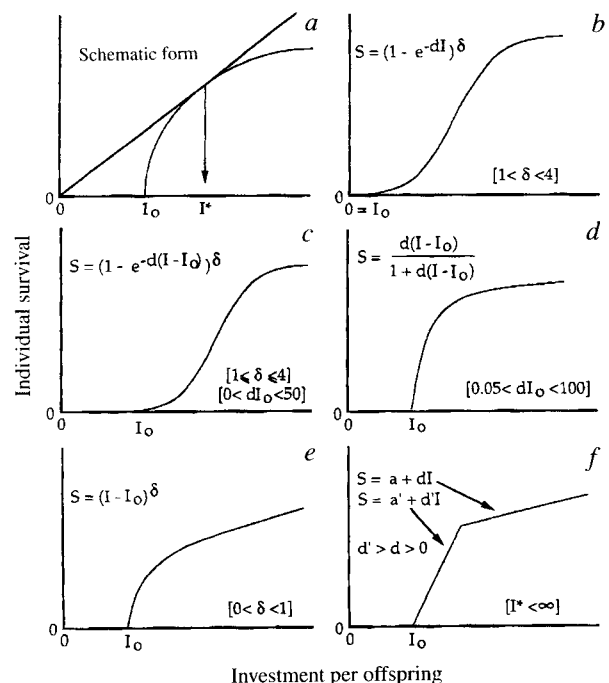
Most evolutionary models for the optimal investment per offspring⁷⁻¹² begin with a trade-off structure (Fig. 1a). Individual offspring survival to adulthood (S) is solely a function of the amount of resource given to the offspring (I). A female with an amount, R , of reproductive resource can produce R/I offspring in a litter, each of which survives with probability $S(I)$. Thus,

the parent will produce $(R/I) S(I)$ surviving offspring. Provided that the parent controls I , natural selection ought to favour the maximization of RS/I through choice of I (called I^* at the maximum); this solution also assumes that the parent's own survival and growth are uncorrelated with I (although they may be related to R). For fixed R , this is equivalent to maximizing S/I ; this optimal I satisfies $\partial S/\partial I = S/I$ and may be found by the tangent argument illustrated in Fig. 1a. We will refer to I as 'offspring size' which is taken to be proportional to investment per offspring (but see below).

This answer for the optimal I assumes R is large relative to I^* so that R/I^* is (effectively) an integer, a necessary condition for real litters. When litter sizes are small (1-6), R/I^* may differ significantly from an integer, and a female may gain more by increasing the size of her C offspring than by producing $C+1$ smaller offspring. Thus there is a range of 'optimal' offspring sizes (I) at each litter size⁷. Suppose that the offspring produced by members of a particular population face the trade-off illustrated in Fig. 1a. Each female (i) has R_i resources to divide up; R_i may differ among females but the $S(I)$ curve does not. Provided that a female does not produce offspring of variable size within her litter, she will have R_i/I_i offspring in each litter with a per-offspring survival of $S(I_i)$. If R is great enough to produce one viable offspring (above I_0 on Fig. 1a), $C=1$ and $I=R$. As R increases, so will I , as long as one offspring of size R will be better than two of size $R/2$. At higher values of R , C will equal 2 and $I=R/2$; still higher values of R yield $C=3$ (and $I=R/3$) and so forth. Each C will be associated with a maximum ($I_{\max}$) and a minimum ($I_{\min}$) offspring size, and resource (R) amounts. The litter size C should change from one to two when the fitness gain from one offspring of size I equals the fitness gain from two offspring of size $I/2$ or $S(I)=2S(I/2)$. The change from two to three offspring should be at size I where $2S(I)=3S(I/3)$. The total resource amount (R_2) at this switch point is the value of R where $2S(R_2/2)=3S(R_2/3)$. When continued to even larger litter sizes, this argument yields successive resource

FIG. 1. Various trade-off functions for offspring survival (S) as a function of investment in each offspring (I). Each equation (in panels a-f) may be multiplied by any positive number without altering any of the results reported here. The trade-off assumed in the classic model^{10,11} is that individual offspring survival to adulthood is solely a function of the resource given to each offspring. A female has some fixed amount of resource (R) to allocate among offspring, so the number of offspring in any litter (C) is R/I . The optimal I maximizes the number of surviving offspring (RS/I). For any fixed R , this is equivalent to choosing I^* to maximize S/I and may be found by the tangent argument illustrated in a. This I^* answer requires that the resulting litter size R/I^* be an integer; this requirement is not well approximated at small litter or clutch sizes (1-6) which are common in many animals. b-f. The invariance/scaling rules discussed in the text hold (to within 5%) for these trade-off curves, over a large range of parameter values (given in brackets) which yield the finite equilibrium I^* in a. Notice that the parameter value ranges refer only to the dimensionless numbers dI_0 and δ ; it is straightforward to show that the rules discussed here only depend on these dimensionless numbers (see below).

METHODS. Let R_i be the total resource level at the change-over from litter size i to $i+1$. Then R_1 is where $S(R_1)=2S(R_1/2)$, R_2 is where $2S(R_2/2)=3S(R_2/3)$, and so forth. Let δ = mean 'to within 5%'; then $R_{i+2}-R_{i+1}=R_{i+1}-R_i$. Because the dimensionless ratio between successive R intervals is only dependent upon the dimensionless numbers in each parametric form (dI_0, δ), calculations are greatly simplified. Dividing by the respective litter sizes yields rule (1). The resource interval ($R_{i+1}-R_i$) is the constant of proportionality (equation (2) in text). Some parametric forms lead to strict equality (for example, f, e with $\delta=0.5$); other show convergence to equality for large i with the '5% or better' rule applicable for small (2-6) litters. The convergence of $R_{i+1}-R_i$ to a fixed value at large i may be developed as follows. The largest (smallest) offspring size is converging to I^* (Fig. 2); for large i we have $R_i/i \approx I^*$ and $(R_{i+1})/(i+1) \approx I^*$. This immediately implies $R_{i+1}-R_i \approx I^*$. Interestingly this is also approximately true for R values yielding clutch sizes of 2-6, although the error here may be up to 10%. Curves e and f are



non-asymptotic, but may usefully represent S over a reasonable I range, particularly as the calculated answers are independent of any positive multiplier of the S function.

amounts (R) at the switch points, and the rule that the largest offspring associated with litters of size C ought to be $(C+1)/C$ times as large as the smallest offspring associated with the next largest litter size, $C+1$. The maximum and minimum offspring sizes converge from above and below to the optimal size I^* , as given in the classic model^{9,10} (Fig. 1a). When offspring size is measured relative to I^* the convergence is then to 1.0 (Fig. 2). We note two limitations to this argument. First, if the parent can easily control R it may well pay simply to accumulate resource until an (or a few) I^* -sized offspring can be produced. Second, the argument assumes that the conversion of R into offspring is linear; significant nonlinearity would distort the observed sizes (investment per offspring) at the switch points. Even given these caveats (R fixed, proportional conversion), the rules discussed here should be broadly applicable.

Our previous work⁷ suggested the theoretical convergence to be more or less symmetric around I^* ; new calculations have proved this to be often not true, but these calculations showed the convergence to have one, unexpected, universal feature; the size range is proportional to C^{-1} (Fig. 2). Except for C_1 , the ratio of the ranges of offspring size at litter sizes i and j are inversely proportional to the ratio of the litter sizes (C_i, C_j); or

$$(I_{\max_i} - I_{\min_i}) / (I_{\max_j} - I_{\min_j}) = C_j / C_i \quad (1)$$

Thus the size range for C_3 is predicted to be 2/3 that at C_2 , and so on; the rule often fails (up to 20–35% off the mark) at C_1 , although it works even here for the parameter values of Fig. 2. This rule is correct to within 5% (often less, even exact) for $2 < C < 6$ for a remarkable variety of trade-off functions, illustrated in Fig. 1. Whereas specific functions and parameter values yield specific values of I^* , $I_{\max_i}$, $I_{\min_i}$, and proportionality constant, each $S(I)$ combined with various and increasing R values yields rule (1). It seems to follow from the economic rule of when a female should switch from a litter size of C to $C+1$ with consequently smaller offspring (compare discussion above and Fig. 1 legend). Rule (1) holds mathematically even for large C values, but is biologically less plausible here.

Writing rule (1) as $\text{OSR} \propto C^{-1}$ (where OSR is offspring size range) implies a proportionality constant (that is, $\text{OSR} = H/C$) and H may be developed as follows. The range in resource values at any litter size k may be written as $R_{\max_k} - R_{\min_k}$. The maximum

($I_{\max}$) and minimum ($I_{\min}$) sized offspring at this litter size are given by $R_{\max_k}/C_k = I_{\max_k}$ and $R_{\min_k}/C_k = I_{\min_k}$. Thus we have

$$\text{OSR}_k = H/C_k = (R_{\max_k}/C_k) - (R_{\min_k}/C_k) \quad (2)$$

and see immediately that the proportionality constant H is simply the resource range at any litter size, $H = R_{\max_k} - R_{\min_k}$. Or said slightly differently, the range of resources yielding any single litter size will be the same for all litter sizes.

There is one other interpretation of H . In Fig. 1 legend we show that, as k gets large, $R_{\max_k} - R_{\min_k} \rightarrow I^*$; as the R intervals are not too different from this even at small C , we have to a good approximation, $H \approx I^*$. If offspring size is measured relative to I^* , as in Fig. 2, then $H \approx 1$ and the intervals are equal to C^{-1} . Notice that this means the entire convergence illustrated in Fig. 2 can be reconstructed approximately from any single point (except I_0/I^*). For example, each $I_{\max_i}/I^*$ will be associated with a convergence, independent of trade-off details. Notice also that this allows an estimate of I^* from two boundary points; that is, $I^* = 2$ (OSR at C_2), and so forth.

Our previous work⁷ estimated the largest egg size (the 90th percentile in volume of a single egg) for litters of 1–5 in the poeciliid fish *Gambusia hubbsi*, and showed a close correspondence to the decline predicted in Fig. 2. This decline is predicted to hold, as a good approximation, for models (Fig. 1) yielding $I_{\max_i}/I^*$ near 1.5. Unfortunately, it proved impossible to estimate the smallest egg size at each litter size so we were (are) unable to estimate the ranges in egg size. We are not aware of any existing data that would allow such estimates in other species.

It has proven exceedingly difficult to get detailed information about the shapes of life-history trade-offs¹¹, and two decades of research (beginning with refs 9 and 10) have yielded very few examples^{8,12} for the estimation of the shape of the trade-off illustrated in Fig. 1. This makes theoretical predictions that are independent of the exact form of trade-offs, like the ones developed here, of particular interest. It is surprising that the numerical prediction of -1 scaling holds for such a large and diverse class of $S(I)$ trade-offs (Fig. 1). This universality makes empirical studies of offspring size ratios especially worthwhile. Size/number trade-offs characterize many other evolutionary problems with similar economic structure¹³, and these may be expected to show similar scaling rules for the size ratios. In addition, the theoretical existence of trade-off invariance for this problem should point theorists to search for numeric universality in other life-history evolution problems⁶. Finally (and empirically), scaling exponents of ± 1 are ubiquitous between timing (size) variables for life histories, and for inter-species (population) comparisons^{6,14–16}; the proportionality constant differs between higher taxa (for example, mammals⁶ versus reptiles¹⁶) but, again, not the exponent. These seem sensible candidates for trade-off-invariance rules. □

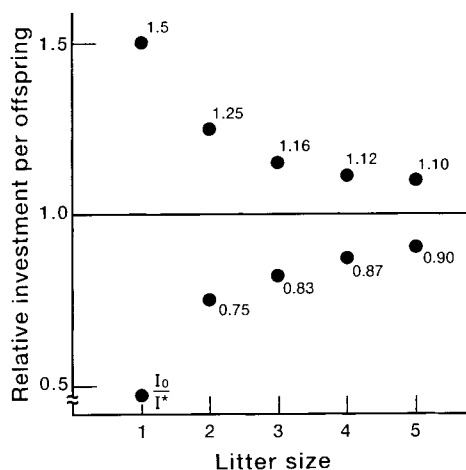


FIG. 2 Maximum and minimum offspring sizes, measured relative to I^* , at different litter sizes for the parametric model $S = (1 - e^{-d(I-I_0)})$ (Fig. 1c with $\delta = 1$). The maximum and minimum values converge on 1.0 from above and below. Here d and I_0 were chosen to make $I_{\max_1}/I^* \approx 1.5$ and the convergence is nearly symmetric. Note that the range of offspring sizes at any litter size (C) is C^{-1} , hence the ratio of ranges at any two litter sizes is the inverse of the ratio of the litter sizes. This scaling rule holds true for a wide range of values of d and I_0 in this model, and for many shapes for the $S(I)$ curve (Fig. 1).

Received 12 April; accepted 14 June 1995.

1. Parker, G. A. & Maynard Smith, J. *Nature* **348**, 27–33 (1990).
2. Charnov, E. L. *The Theory of Sex Allocation* (Princeton Univ. Press, 1982).
3. Stephens, D. W. & Krebs, J. R. *Foraging Theory* (Princeton Univ. Press, 1986).
4. Bulmer, M. G. *Theoretical Evolutionary Ecology* (Sinauer, Sunderland, MA, 1994).
5. Godfray, H. C. J. *Parasitoids* (Princeton Univ. Press, 1994).
6. Charnov, E. L. *Life History Invariants* (Oxford Univ. Press, 1993).
7. Charnov, E. L., Downhower, J. F. & L. Brown, *Evol. Ecol.* **9**, 57–63 (1995).
8. Clutton Brock, T. *The Evolution of Parental Care* (Princeton Univ. Press, 1991).
9. Schaffer, W. M. & Gadgil, M. in *Ecology and Evolution of Communities* (eds Cody, M. & Diamond, J. M.) 142–157 (Harvard Univ. Press, Cambridge, MA, 1975).
10. Smith, C. C. & Fretwell, S. D. *Am. Nat.* **108**, 499–506 (1974).
11. Stearns, S. C. *The Evolution of Life Histories* (Oxford Univ. Press, 1992).
12. Wootton, R. J. *J. Fish. Biol.* **45**, 1067–1077 (1994).
13. Lloyd, D. G. *Am. Nat.* **129**, 800–817 (1987).
14. Charnov, E. L. & Berrigan, D. *Evol. Ecol.* **4**, 273–275 (1990).
15. Beverton, R. J. H. *J. Fish. Biol.* **41**(s), 137–160 (1992).
16. Shine, R. & Iverson, J. B. *Oikos* **72**, 343–348 (1995).

ACKNOWLEDGEMENTS. We thank S. Proulx for extensive computer calculations and discussions, and J. J. Bull for improving the text and title. The Ministry of Agriculture and Fisheries of the Bahamas provided permission to study *G. hubbsi* populations. This work was supported by the Biology Department of the University of Utah, Earthwatch and the National Geographic Society.

Hox genes and the diversification of insect and crustacean body plans

Michaëls Averof & Michael Akam

Wellcome/CRC Institute, Tennis Court Road, Cambridge CB2 1QR, and Department of Genetics, University of Cambridge, Cambridge, UK

CRUSTACEANS and insects share a common origin of segmentation^{1,2}, but the specialization of trunk segments appears to have arisen independently in insects and various crustacean subgroups^{3,4}. Such macroevolutionary changes in body architecture may be investigated by comparative studies of conserved genetic markers⁵⁻⁷. The *Hox* genes are well suited for this purpose, as they determine positional identity along the body axis in a wide range of animals⁸⁻¹⁰. Here we examine the expression of four *Hox* genes in the branchiopod crustacean *Artemia franciscana*, and compare this with *Hox* expression patterns from insects. In *Artemia* the three 'trunk' genes *Antp*, *Ubx* and *abdA* are expressed in largely overlapping domains in the uniform thoracic region, whereas in insects they specify distinct segment types within the thorax and abdomen. Our comparisons suggest a multistep process for the diversification of these *Hox* gene functions, involving early differences in tissue specificity and the later acquisition of a role in defining segmental differences within the trunk. We propose that the branchiopod thorax may be homologous to the entire pregenital (thoracic and abdominal) region of the insect trunk.

We have previously shown that gene duplication and diversification in the arthropod *Hox* clusters pre-date the divergence of crustaceans and insects^{5,11}. Thus, in spite of striking differences in body organization, crustaceans and insects share similar sets of homeotic genes. Here, we use *in situ* hybridization and immunohistochemical techniques to study expression of several of these genes during the larval (postnaupliar) stages of *Artemia* development, while trunk segments are forming^{12,13} (Fig. 1).

The monoclonal antibody FP6.87 recognizes a conserved epitope in the *Ubx* and *AbdA* proteins of diverse insects^{14,15}. This epitope has been mapped to a region which is also conserved in the *Ubx* and *AbdA* proteins of *Artemia* (Fig. 2a); we therefore use this antibody to examine *Ubx* and *abdA* expression during *Artemia* development (Fig. 2d-h). We observe staining with nuclear localization restricted to the thoracic region of the body. Staining first appears when the thoracic region is forming, before segments are externally distinct (Fig. 2d). The expression has a discrete anterior boundary that persists throughout development and co-incides with the morphological boundary between the gnathal and postgnathal regions of the body. Ventrally, this boundary appears to be parasegmental^{1,2}, starting in the posterior part of the last gnathal segment (Mx2; Fig. 2f). In the body wall, staining starts from the first thoracic segment (T1) and extends backwards during development, as new segments are formed from the posterior growth zone (Fig. 2g, h). Expression extends transiently into the prospective genital and abdominal segments, but later retracts and is confined to the thorax (as shown for *abdA* in Fig. 2k). As thoracic segments develop (in antero-posterior progression), staining becomes strongest in the developing neuromeres of the central nervous system and then gradually fades. In mature segments levels of expression appear to be low.

To explore the expression of *abdA* and *Ubx* further, we have raised antibodies against a peptide which is diagnostic for the *AbdA* protein and is conserved between *Artemia* and insects (Fig. 2a). In insects this antibody can specifically detect the *AbdA* protein distribution (Fig. 2b, c). We can therefore use this antibody to examine *abdA* expression in *Artemia* independently from *Ubx*. Unfortunately, owing to the low sensitivity of the

antibody, we cannot detect expression during early stages, but we obtain clear nuclear staining during middle to late stages of postnaupliar development (Fig. 2i-k). This confirms that the *AbdA* protein is expressed in the *Artemia* thorax but is absent in the abdomen. Expression appears strongest in the neuromeres and weaker in the body wall and limbs. The anterior boundary of expression during early stages is unclear, but at later stages we detect weak expression in a few cells of the T2 neuromere and in many cells of the T3 neuromere, and stronger widespread expression in the more posterior neuromeres of the thorax (Fig. 2i, j). A few cells stain in the genital segments (Fig. 2k), but not in genital neuromeres.

In situ hybridization was used to examine the expression of *Antp* and *AbdB*. *Antp* is expressed throughout the *Artemia* thorax and extends anteriorly into the last gnathal segments (posterior Mx1 and entire Mx2; Fig. 2l). Expression appears strongest in the limbs, but is weak or absent in the neuromeres, where *abdA* and perhaps *Ubx* are expressed at similar developmental stages. *AbdB* is expressed in the genital segments of *Artemia*, immediately posterior to the domain where *Antp*, *Ubx* and *abdA* are coexpressed (Fig. 2m).

The expression patterns described are consistent with a role of these genes in the positional specification of segmental fate; they are limited to well-defined regions of the trunk, and appear before visible segment differentiation. Our results suggest that *Antp*, *Ubx* and *abdA* are collectively involved in defining the segmental identity of the thoracic region. Although these genes may have different spatio-temporal or cell-specific functions in the thorax (such as late function of *abdA* predominantly in the neuromeres), they appear not to be used to define distinct segment identities within the postgnathal trunk; all thoracic segments have identical morphology, except for small differences in size (Fig. 1). *Antp* may also be involved in the specification of the posterior gnathal region, although we would predict on the basis of conventional homologies¹⁶ that *Scr* would be more important in defining segmental identity in this region¹⁷.

The expression of *Antp*, *Ubx* and *abdA* in *Artemia* contrasts with what has been described in insects, where these genes are expressed in discrete (although still partly overlapping) domains along the trunk and are responsible for specifying the differential fates of thoracic and abdominal segments^{8,18,19}. Most striking is the expression of *abdA* in the *Artemia* thorax; in insects it is expressed exclusively in the abdomen, and its function is thought to define the identity of that region^{15,19-21}. We consider the contrasting expression patterns of these genes in *Artemia* and insects under a general model for the evolution of *Hox* gene functions (Fig. 3).

The regulation of eukaryotic genes depends on distinct regulatory elements, each responsible for different spatio-temporal or tissue-specific aspects of a complex pattern. Thus we expect that when ancestral *Hox* genes duplicate, different subsets of these regulatory elements may become partitioned into separate genes. If so, these duplications would produce functionally identical homeoproteins expressed in different tissues, regions or developmental stages. Each of the newly duplicated genes would then be subject to independent selection on the basis of its unique features of expression. Similar cases of regulatory diversification have been documented for duplicated genes with functionally identical protein products^{22,23}. We speculate that such a process occurred in some common ancestor of crustaceans and insects, giving rise to the closely related *Antp*, *Ubx* and *abdA* genes; initially these were involved in specifying a common regional identity, the 'trunk', with no distinct segment-specific functions. Furthermore, we suggest that this ancestral state may be reflected in the overlapping expression domains of *Antp*, *Ubx* and *abdA* in *Artemia*, and to a large extent of *Ubx* and *abdA* in *Drosophila*²⁰. Even today, *AbdA* and *Ubx* proteins have identical effects in many genetic assays^{24,25}. An analogous situation has also been documented in vertebrates, where duplications of

entire *Hox* gene clusters gave rise to paralogous sets of genes with partly redundant functions in the hindbrain¹⁰.

In the lineage that gave rise to the insects it appears that the expression of these genes resolved to more discrete antero-posterior domains, through changes in gene regulation and competing cross-regulatory interactions¹⁰. Ultimately, the diversification of *Hox* gene functions to regulate different sets of downstream genes within each of these domains led to the morphological diversification of segments in the thorax and abdomen of insects. One example would be the repression of *Distal-less* expression by *Ubx* and *abdA*, and the consequent development of a limbless abdomen in insects^{15,25}.

Our observations provide a framework for direct morphological comparisons between arthropod body plans, where positional homologies can be defined on the basis of underlying

genetic information (Fig. 4a). We suggest that the thoracic region of branchiopods is homologous to the entire pregenital trunk region of the insects (thorax and most of the abdomen), a region that we designate as 'trunk'. It is intriguing that the number of segments in this region is conserved between most branchiopods and insects, with only limited variations in each group. This scheme identifies the genital segments as homologous in these two groups, defined by the activity of *AbdB*²⁶, and occupying conserved positions at the end of the trunk. It also implies that the *Artemia* abdomen is a terminal structure, perhaps comparable to the highly modified and reduced postgenital segments of the insect abdomen.

The proposed regional homologies will have to be reconciled with the phylogenetic status and body plans of other arthropod groups. At present, the proposed scheme may be extended to

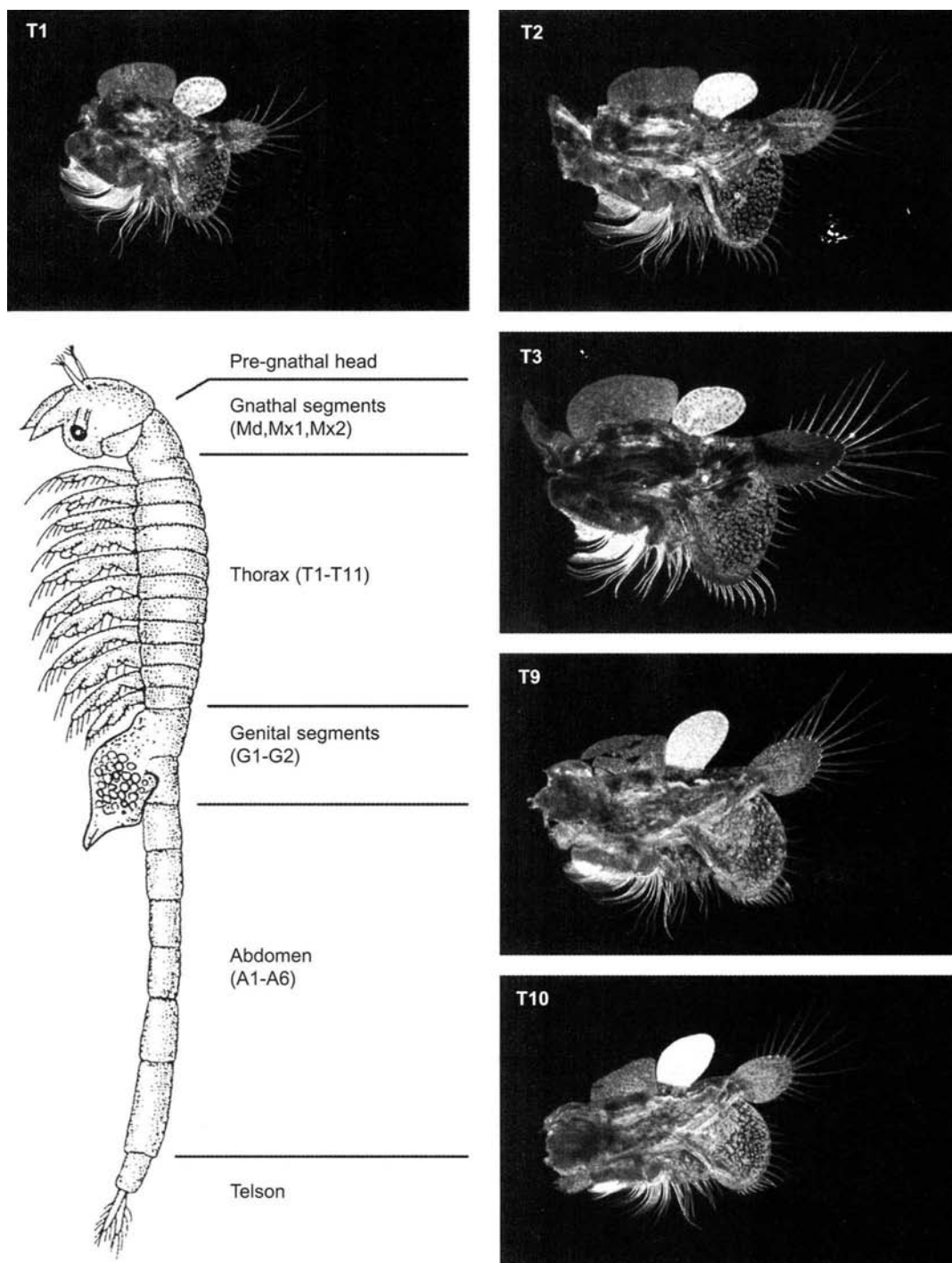


FIG. 1 A diagrammatic representation of the *Artemia* body plan, consisting of a head (including 3 gnathal segments; Md, Mx1 and Mx2), thorax (11 segments: T1-T11), genital region (2 segments: G1 and G2), abdomen (6 segments: A1-A6) and telson. Apart from differences in size, all thoracic segments have similar morphology. We illustrate this by showing the structure of the limbs on thoracic segments T1, T2, T3, T9 and T10.

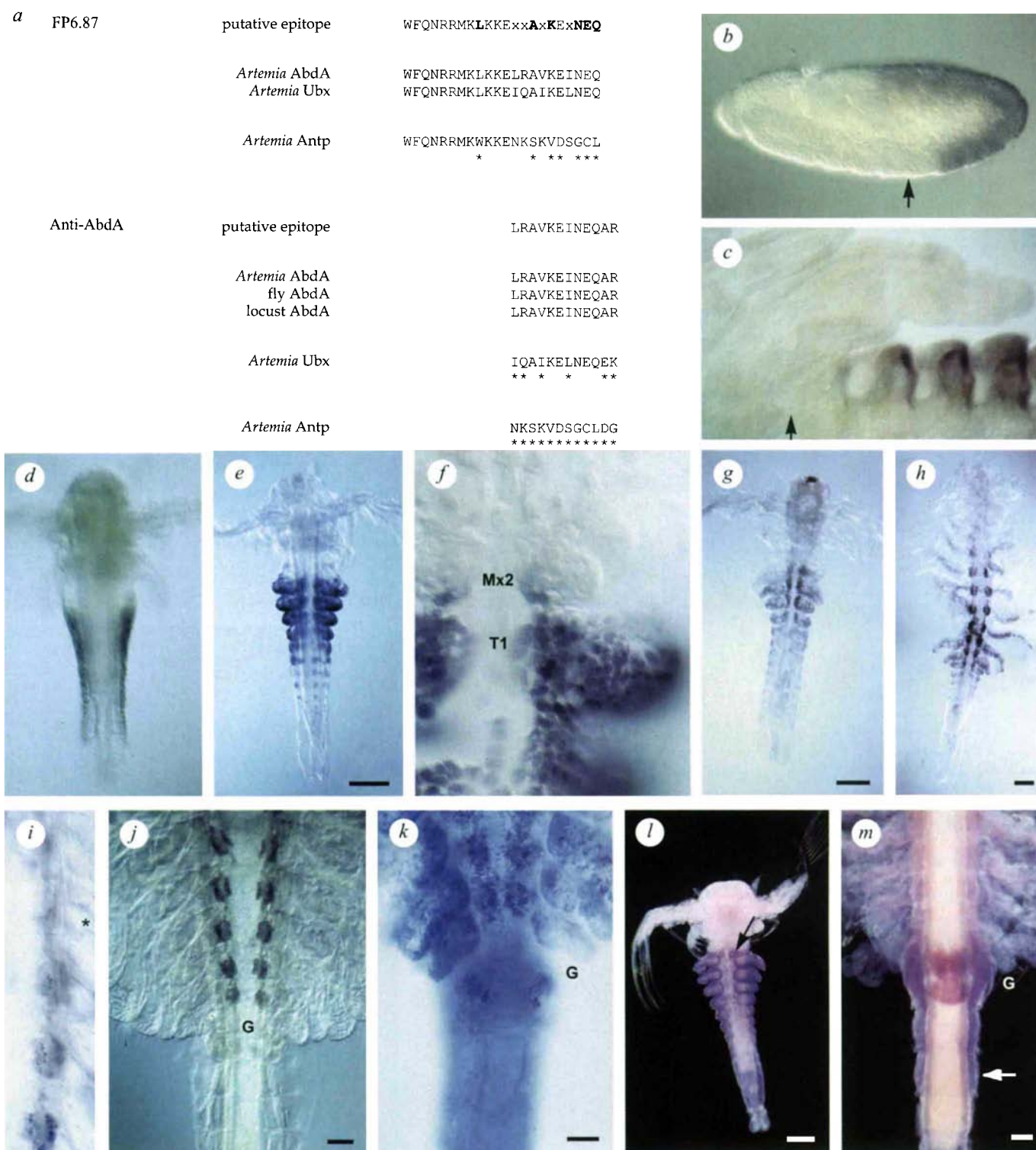


FIG. 2 *Hox* gene expression in *Artemia franciscana*. **a**, The putative epitopes recognized by the FP6.87 (ref. 14) and anti-AbdA antibodies are compared to the protein sequence of AbdA, Ubx and Antp from *Artemia franciscana*⁵. Both epitopes lie near the carboxy-terminal end of the homeodomain. Residues in bold are thought to define the specificity of the FP6.87 antibody; asterisks indicate differences between the putative epitope and a particular homeoprotein. **b**, **c**, Specificity of the anti-AbdA serum. Immunohistochemical staining on fly (**b**) and locust (**c**) embryos using anti-AbdA reveals a pattern identical to the previously characterized AbdA protein distribution^{20,21} but with low sensitivity (arrowheads indicate the borders between abdomen and thorax). **d**–**h**, Staining with the FP6.87 antibody during *Artemia* development. Staining appears very early, before external segmentation of the thorax (**d**), and persists throughout development of the thoracic region (**e**, **g**, **h**) in a domain that starts at the posterior end of the Mx2 limb bud (**f**) and spans the entire thorax. **i**–**k**, Staining with the anti-AbdA antibody in *Artemia*. At relatively late stages, strong staining can be seen in the thoracic neuromeres, starting weakly in T2 (marked by asterisk in **i**) and extending posteriorly throughout the thorax. At late stages,

staining appears in a few cells within the genital region (**G**) (**k**), but is completely absent from the genital neuromeres and abdomen (**j**, **k**). **l**, Antp expression in *Artemia*. *In situ* hybridization shows expression in the developing thorax, in the limb buds of Mx2 (indicated by the arrowhead) and the posterior of Mx1. **m**, AbdB expression in *Artemia*. *In situ* hybridization with a short probe shows expression in the two genital segments (**G**) and the gonads. The gonads develop in the genital region and extend laterally into the abdomen (arrowhead). Scale bars, 0.1 mm.

METHODS. The anti-AbdA antiserum was raised against a synthetic AbdA-specific peptide (LRAVKEINEQAR). This was coupled to ovalbumin through an amino-terminal cysteine and used to immunize rabbits. The serum was affinity-purified using a locust AbdA fusion protein²¹. Immunohistochemical stainings were performed by standard methods¹⁴, after sonication¹³. *In situ* hybridizations were performed with digoxigenin-labelled antisense RNA probes for *AfAntp*⁵ and *AfAbdB* (method adapted from ref. 13). The *Artemia* *AbdB* fragment was isolated by using degenerate-primer PCR (F. Casares, unpublished data; sequence, EMBL accession number X87250).

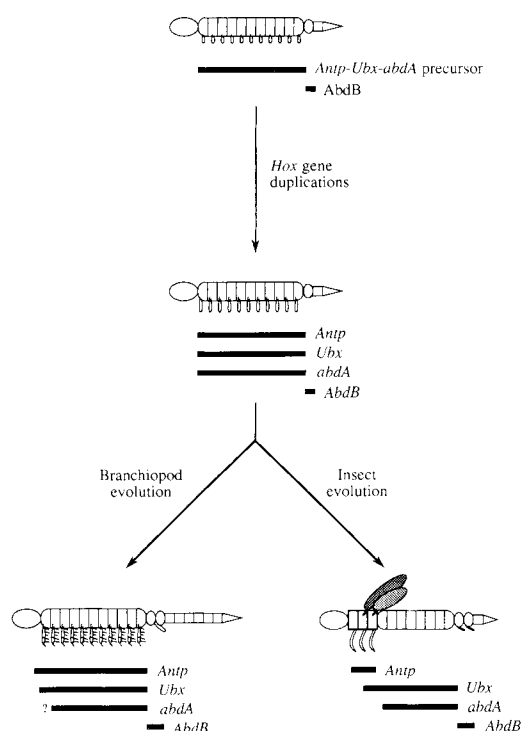


FIG. 3 A model for the evolution of *Hox* gene functions. We expect that the progenitor of *Antp*, *Ubx* and *abdA* was a *Hox* gene responsible for specification of the middle body region (the 'trunk') in a distant proto-stome ancestor. The gene duplications that gave rise to distinct *Antp*, *Ubx* and *abdA* genes must have occurred before the insect and crustacean lineages diverged^{5,11}. These duplications, producing genes with functionally identical protein products, allowed the duplicated copies to acquire slightly different temporal, spatial, tissue- or cell-specific regulation. Further elaboration of this regulation (resolution of spatial domains of expression) and acquisition of differential downstream functions led to diversification of this region into the thorax and abdomen of insects. We expect that *AbdB* acquired a distinct function in the 'tail' region much earlier (before the divergence of major animal phyla), and that at some point during arthropod evolution it became associated with the genital segments.

include the body plans of phylogenetically interesting fossils such as *Rehbachella*, *Lepidocaris* and the euthycarcinoids^{4,27-29} (Fig. 4b). Interestingly, euthycarcinoids were originally identified as branchiopod crustaceans²⁸, and more recently were thought to be associated with the origin of the insects²⁹. More problematic is the position of the myriapods and other crustacean classes, but plausible homologies which are consistent with our scheme have been suggested^{4,30} (Fig. 4b). We expect that many of these issues will be resolved by extending our present study to diverse arthropod groups. □

Received 28 March; accepted 2 June 1995.

- Dohle, W. & Scholtz, G. *Development* **104** (suppl.), 147-160 (1988).
- Patel, N. H., Kornberg, T. B. & Goodman, C. S. *Development* **107**, 201-212 (1989).
- Calman, W. T. *Crustacea* (Black, London, 1909).
- Wolossek, D. *Fossils and Strata* **32**, 1-202 (1993).
- Averof, M. & Akam, M. *Curr. Biol.* **3**, 73-78 (1993).
- Slack, J. M. W., Holland, P. W. H. & Graham, C. F. *Nature* **361**, 490-492 (1993).
- Arendt, D. & Nubier-Jung, K. *Nature* **371**, 26 (1994).
- Lewis, E. B. *Nature* **276**, 567-570 (1978).
- Akam, M., Dawson, I. & Tear, G. *Development* **104** (suppl.), 123-133 (1988).
- McGinnis, W. & Krumlauf, R. *Cell* **68**, 283-302 (1992).
- Akam, M. et al. *Development* **120** (suppl.), 209-215 (1994).
- Anderson, D. T. *Aust. J. Zool.* **15**, 47-91 (1967).
- Manzanares, M., Marco, R. & Garesse, R. *Development* **118**, 1209-1219 (1993).
- Kelsh, R., Weinzierl, R. O. J., White, R. A. H. & Akam, M. *Dev. Genet.* **15**, 19-31 (1994).
- Warren, R. W., Nagy, L., Selegue, J., Bates, J. & Carroll, S. *Nature* **372**, 458-461 (1994).
- Snodgrass, R. E. *Comparative Studies on the Head of Mandibulate Arthropods* (Comstock, New York, 1951).
- Mahaffey, J. W., Diederich, R. J. & Kaufman, T. C. *Development* **105**, 167-174 (1989).

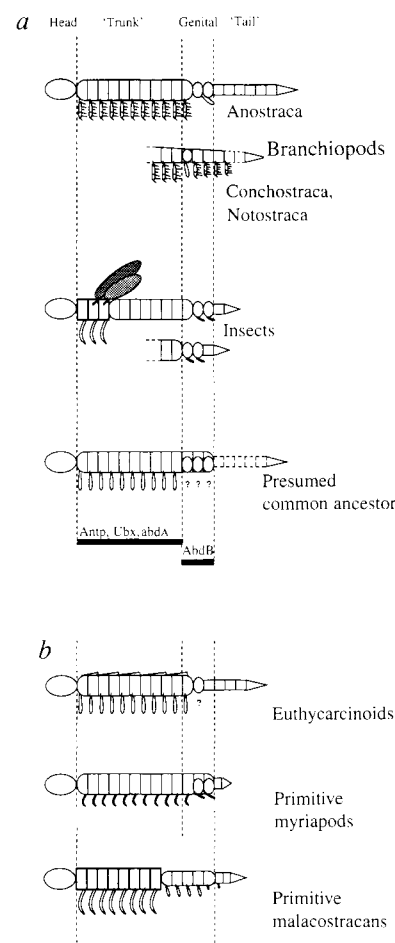


FIG. 4 a, A hypothesis for regional homologies between branchiopod and insect body plans. The homology of individual segments in the head region of crustaceans and insects (including the gnathal segments) is widely accepted¹⁶. We suggest that the thoracic region of branchiopod crustaceans (including Anostraca, Conchostraca and Notostraca) is homologous to the entire pregenital (thoracic and abdominal) region of the insects, and that genital segments hold homologous positions in both these groups. Genital segments are indicated by an oval shape. Our hypothesis suggests that the last common ancestor of crustaceans and insects was an animal with a distinct mandibulate head, a trunk defined by the activity of *Antp*, *Ubx* and *abdA*, genital segments defined by the activity of *AbdB*, and a postgenital 'tail'. b, Suggested homologies to other arthropod body plans. The body plans of euthycarcinoids^{28,29} can be accommodated in our scheme. Homologies to myriapods and malacostracan crustaceans are more problematic, but the suggestions of Dohle³⁰, on the nature of primitive myriapods, and Walossek⁴, on the homology between the branchiopod thorax and most of the malacostracan trunk (abdomen and thorax), provide some plausible solutions.

- Beeman, R. W., Stuart, J. J., Brown, S. J. & Denell, R. E. *BioEssays* **15**, 439-444 (1993).
- Ueno, K., Hui, C. C., Fukuta, M. & Suzuki, Y. *Development* **114**, 555-563 (1992).
- Karch, F., Bender, W. & Weiffenbach, B. *Genes Dev.* **4**, 1573-1588 (1990).
- Tear, G., Akam, M. & Martinez-Arias, A. *Development* **110**, 915-925 (1990).
- Li, X. & Noll, M. *Nature* **367**, 83-87 (1994).
- Rothe, M., Pehl, M., Taubert, H. & Jackie, H. *Nature* **359**, 156-159 (1992).
- Castelli-Gair, J., Greig, S., Micklem, G. & Akam, M. *Development* **120**, 1983-1995 (1994).
- Vachon, G. et al. *Cell* **71**, 437-450 (1992).
- Jurgens, G. & Hartenstein, V. in *The Development of Drosophila melanogaster* (eds Bate, M. & Martinez-Arias, A.) 687-746 (Cold Spring Harbor Laboratory Press, 1993).
- Scourfield, D. J. *Phil. Trans. R. Soc. B* **214**, 153-187 (1926).
- Riek, E. F. *Rec. S. Aust. Mus.* **27**, 313-321 (1968).
- McNamara, K. J. & Trewin, N. H. *Palaeontology* **36**, 319-335 (1993).
- Dohle, W. *Myriapoda and the Ancestry of Insects* (Manchester Polytechnic, Manchester, 1988).

ACKNOWLEDGEMENTS. We thank M. Manzanarez and L. Nagy for sharing their results on antibody stainings; F. Casares for isolating the *AbdB* PCR fragment; N. Patel, J. Castelli-Gair and F. Ferrari for discussion; and A. Mitchell for help with *in situ* hybridization. This work has been supported by the Wellcome Trust.

Distinguishable functions for *engrailed* and *invected* in anterior-posterior patterning in the *Drosophila* wing

Andrew J. Simmonds*, William J. Brook†, Stephen M. Cohen† & John B. Bell*‡

* Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

† European Molecular Biology Laboratory, Meyerhofstrasse 1, D69117 Heidelberg, Germany

‡ To whom correspondence should be addressed

SUBDIVISION of the limb primordia into compartments initiates pattern formation in the developing limbs^{1,2}. Interaction between distinctly specific cells in adjacent compartments leads to localized expression of the secreted signalling molecules Wingless (Wg) or Decapentaplegic (Dpp), which in turn organize pattern and control growth of the limbs³⁻⁸. The homeobox gene *engrailed* has been implicated in specification of posterior cell fate⁸⁻¹², whereas the LIM/homeobox gene, *apterous*, specifies dorsal fate³. Removing *apterous* activity causes a complete transformation from dorsal to ventral fate and leads to the formation of an ectopic dorsal-ventral boundary organizer^{3,13}. By contrast, removing *engrailed* activity causes incomplete morphological transformation from posterior to anterior fate in the wing^{10,14,15}, and fails to produce an ectopic anterior-posterior organizer (reviewed in ref. 2). Complete transformation can only be effected by simultaneously eliminating activity of *engrailed* and its homologue *invected*¹⁶⁻¹⁸. Here we show that *invected* functions principally to specify posterior cell fate. Thus establishment of the anterior-posterior organizer and control of compartment identity are genetically distinguishable, and *invected* may perform a discrete subset of functions previously ascribed to *engrailed*.

A dominant mutation at the *invected* (*inv*) locus was identified on the basis of a phenotypic transformation of the anterior compartment of the wing into a posterior compartment. The mutant produces wings with a symmetric double-posterior appearance

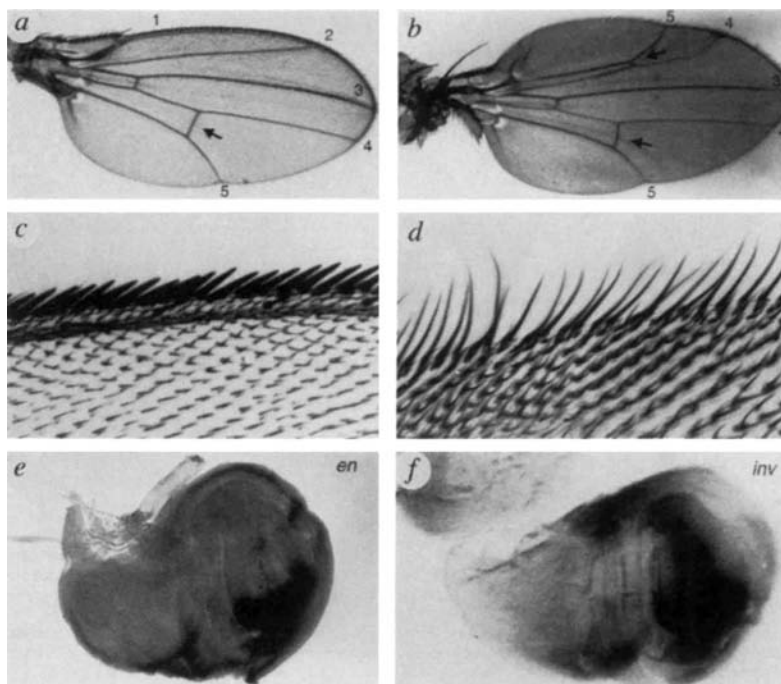
(Fig. 1a, b). Transformation from anterior to posterior cell fate is complete at the wing margin (Fig. 1c, d) and through much of the wing blade (Fig. 1b). Molecular and cytological analyses showed that the mutation is the result of a chromosomal inversion affecting the *invected* and *vestigial* (*vg*) genes. The juxtaposition of *vg* and *inv* sequences in the inversion chromosome leads to ectopic expression of *inv* messenger RNA in the anterior compartment (Fig. 1f), but does not affect expression of *engrailed* (*en*) mRNA, which remains posteriorly localized (Fig. 1e). *Invected* protein is misexpressed in a broad stripe straddling the dorsal-ventral (D-V) boundary in the dominant mutant wing disc, overlapping the expression domain of the endogenous *vestigial* protein (Fig. 2a, b). Thus it appears that *inv* coding sequences have been brought under the control of *vg* enhancer elements as a result of the inversion. We refer to this mutation as *inv*^{Dominant} (*inv*^D). The *inv*^D phenotype is reciprocal to the symmetric double anterior wing produced by the *en*¹ mutation^{9,19}.

Although cell fates are transformed in the anterior compartment, *inv*^D wings are symmetric with respect to the anterior-posterior (A-P) compartment boundary. The axis of symmetry of the wing is defined by localized expression of the secreted signalling molecule Dpp in a row of anterior cells near the compartment boundary^{4,7,8,20}. Despite the level of ectopic *invected* protein in the *inv*^D wing disc being sufficient to transform the morphological identity of anterior cells, it does not cause ectopic expression of *hedgehog* (*hh*) or *dpp* in the anterior compartment (Fig. 2c-f). Ectopic *inv* protein has little or no effect on the patterns of *patched* (*ptc*) or *cubitus interruptus* (*ci*) expression in the anterior compartment, at most causing a slight reduction in the levels of expression near the D-V boundary (Fig. 2e-h). The *inv* protein can therefore direct cells to adopt posterior identity without affecting the expression of genes, such as *hh* and *dpp*, which mediate the patterning activities of the compartment boundary.

By contrast, ectopic expression of *en* protein in the anterior compartment causes pattern abnormalities distinct from those caused by ectopic *Inv* (Fig. 3). Expression of *En* in the anterior hinge region causes bifurcation of the anterior compartment (Fig. 3a, b) in a manner that resembles the effects of ectopic *hh* or *dpp* expression^{4,7,8,20}. These observations suggest that the specification of posterior identity by *Inv* depends on a set of

FIG. 1 Anterior to posterior cell fate transformation due to ectopic expression of *invected* in the anterior compartment. a, Wild-type and b, *inv*^D wings, respectively. The veins serve as markers for position in the anterior and posterior compartments. Veins 1-3 are anterior; veins 4 and 5 are posterior and are connected by the posterior cross-vein (arrow). Veins 1 and 2 are replaced by veins 4 and 5 in the mutant wing. Note the duplicated posterior cross-vein in the mutant wing, and the overall change in the shape of the anterior wing to resemble the posterior compartment. c, Detail of the wild-type anterior wing margin showing the structure of the triple row. d, Detail of the transformed anterior margin of an *inv*^D wing. Note the altered morphology of the bristles, which is typical of the wild-type posterior margin (not shown). e, f, *In situ* hybridization with RNA probes specific for *en* and *inv* transcripts. Note that *inv* transcript is expressed in the anterior compartment, but *en* is not.

METHODS. The mutation *inv*^D was isolated as a revertant of *vestigial*^W, a dominant mutation linking the *vestigial* and *invected* genes via a chromosomal inversion³¹. A 988-bp *EcoRI*-*Bst*XI fragment from the 5' region of the *invected* complementary DNA was used as a template for synthesis of an *invected*-specific RNA probe. A 1,200-bp fragment from the 5' end of an *en* cDNA that has no similarity to *inv* was used as a template for synthesis of an *en*-specific RNA probe.

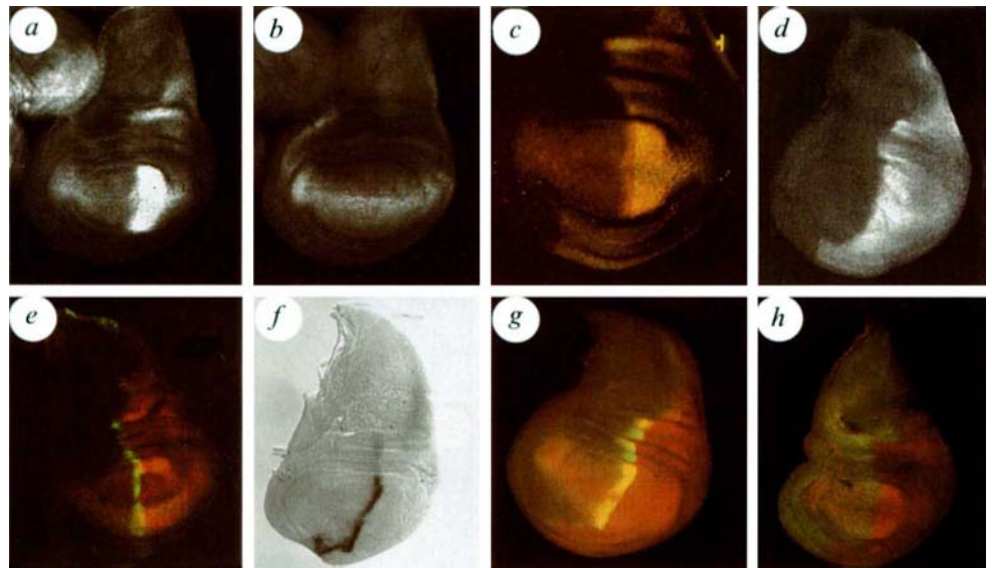


effector genes distinct from those through which En-expressing cells direct formation of the compartment boundary organizer.

To compare the effects of ectopic En and Inv more directly, we misexpressed En in a stripe along the D-V boundary. Under these conditions the pattern of En expression is similar to that

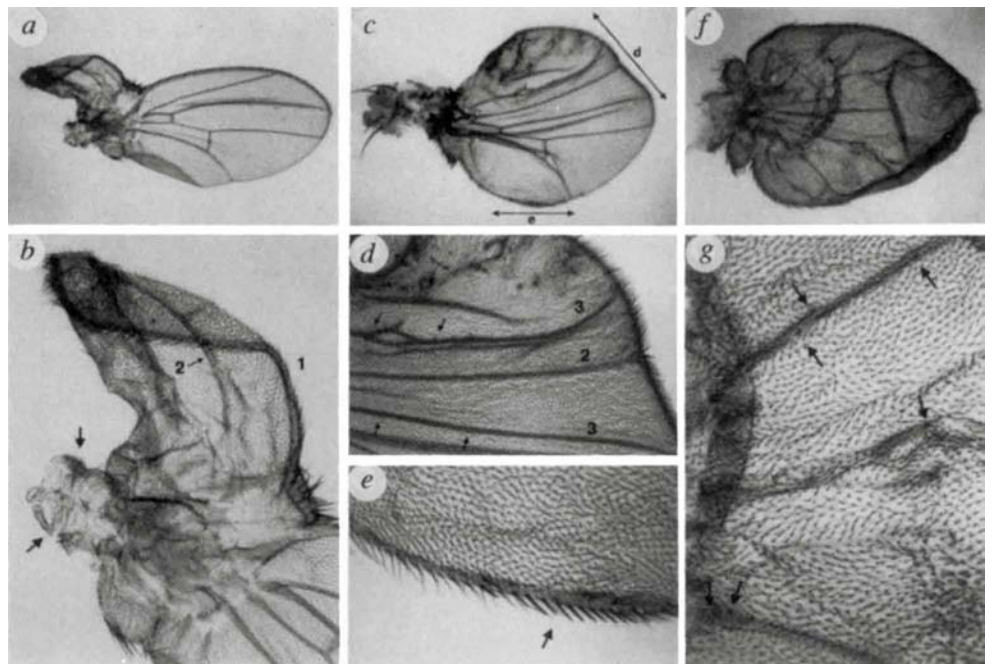
of Inv in *inv^D* mutant wings (compare Fig. 2a with Fig. 4b). In contrast, ectopic En causes pattern duplications in the anterior compartment (Fig. 3c-g). Anteriorly positioned cells adopt the identities of cells normally located near the compartment boundary. Note the presence of a duplicated vein 3 anterior and the

FIG. 2 Ectopic *invected* protein does not affect the expression patterns of *hedgehog*, *decapentaplegic*, *patched* or *cubitus interruptus*. Expression of a, *invected/engrailed* proteins and b, *vestigial* protein in a single *inv^D* wing disc. Although the antibody recognizes both *inv* and *en* proteins, only *inv* transcript is expressed in the anterior compartment (see Fig. 1e, f). The anterior expression of *inv* protein overlaps the spatial domain of *vg* protein expression. c, Double label showing *inv/en* protein (red) and *hh-lacZ* expression (green) in an *inv^D* wing disc. Note that the Hh domain appears yellow (rather than green) owing to overlap with En and Inv. d, Hh is restricted to the posterior compartment in an *inv^D* wing disc. e, Double label showing *inv/en* protein (red) and *dpp-lacZ* expression (green) in an *inv^D* wing disc. f, *dpp* transcript in an *inv^D* wing disc. g, Double label showing *inv/en* protein (red) and *ptc-lacZ* expression (green) in an *inv^D* wing disc. h, Double label showing *inv* protein (red) and *ci* protein expression (green) in an *inv^D* wing disc.



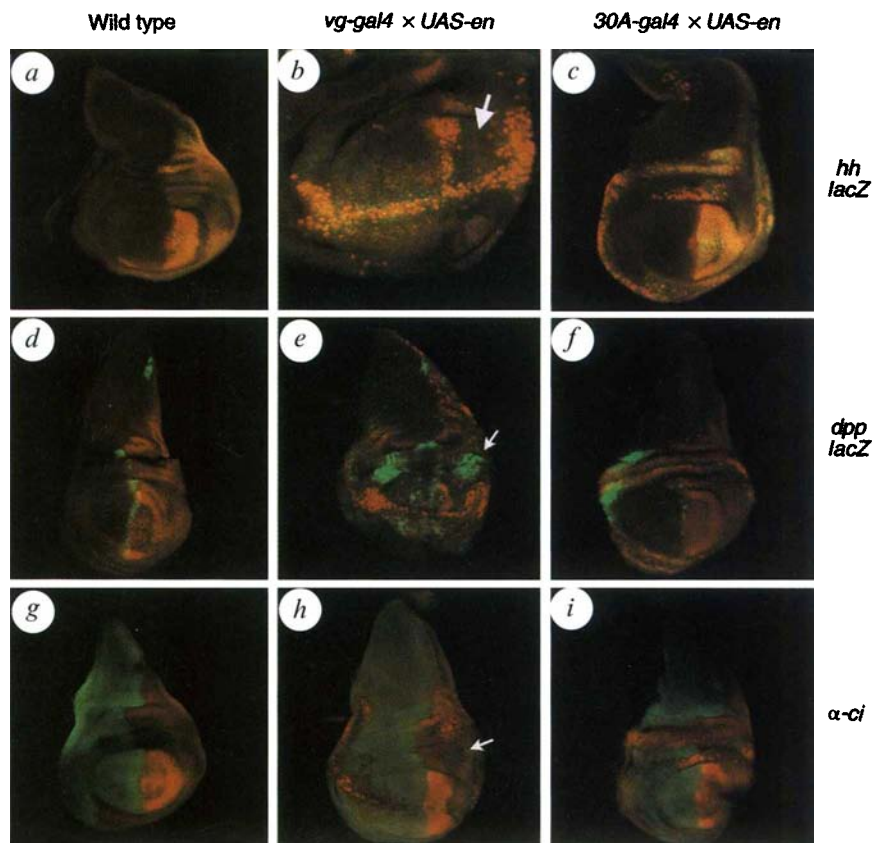
En/Inv²⁴, rabbit anti-Vg²⁵, rabbit anti-βGAL, rabbit anti-Hh⁵ or rabbit anti-Ci (ref. 26, kindly provided before publication by C. Schwartz and T. Kornberg). The *dpp-lacZ*, and *hh-lacZ* reporter genes are described in refs 27, 28; *ptc-lacZ* was a gift from P. Ingham.

FIG. 3 Pattern abnormalities caused by ectopic Engrailed expression in the anterior compartment. a, Axis bifurcation and transformation of the anterior wing hinge caused by ectopic expression of En. The GAL4 driver 30A (ref. 29) was used to direct En expression to the hinge region. b, Detail of the wing shown in a. Note the presence of vein 1 and 2 structures in the supernumerary wing, and the duplication of the axillary cord at the base of the hinge (arrows). c-g, Cell fate transformation and pattern reorganization caused by expression of upstream activator sequence (UAS)-En along the wing margin under the control of a Vg-GAL4 driver. c, Note the partial duplication of anterior compartment pattern elements. d, Detail of the wing shown in c, showing the formation of an ectopic vein 3 anterior to the endogenous vein 2. Vein 3 is identified by its dorsal location and by the presence of campaniform sensillae (arrows). Note that the polarity of the wing reverses at vein 2 and the anterior wing margin bristles are transformed to resemble those normally found between veins 2 and 3 in the anterior compartment. e, Partial transformation of posterior wing margin to anterior identity (arrow). f, Complete anterior to posterior transformation of the wing hinge, together with a 'paradoxical' posterior to anterior transformation of the wing blade. Note the symmetric appearance of the hinge, which contains duplicated axial cord and alula. The wing margin is entirely anterior. g, A broad region of the internal wing blade has adopted the identity of the vein 3 region indicated by campaniform sensillae on all veins.



METHODS. The Vg-GAL4 driver was constructed by S. M. Morimura and F. M. Hoffmann (manuscript submitted). The 30A driver is described in ref. 29. The *en* cDNA clone D2B (ref. 30) was cloned into the pUAST vector²⁹ to construct UAS-*en*. A comparable UAS-*inv* construct was generated. UAS-*inv* is transcribed under control of the GAL4 drivers but does not produce *inv* protein, suggesting that the construct cannot be translated (not shown; similar results were obtained with UAS-*inv* by G. Morata and T. Kornberg (personal communication)). Thus it is not possible to make the optimal direct comparison between *en* and *inv* misexpression using the GAL4 system.

FIG. 4 Ectopic *engrailed* protein alters the patterns of *hedgehog*, *decapentaplegic* and *cubitus interruptus* expression. *a-c*, Double labelling to visualize *en* protein (red) and *hh-lacZ* expression (green). *a*, Wild type. *b*, UAS-En misexpressed using the Vg-GAL4 driver. Note that En is expressed at a high level along the D-V boundary and at lower levels more broadly in the anterior wing blade. *hh-lacZ* is expressed throughout the region where *en* is expressed, independent of the level of *en* expression. Comparable levels of *inv* protein do not activate Hh (compare with Fig. 2c, d). Note that the monoclonal antibody recognizes an epitope common to the homeodomains of *en* and *inv*²⁴, and as such is likely to recognize the two proteins equally. The arrow indicates a patch of cells lacking *en/inv* expression in the posterior compartment. Although *en* can be expressed at high levels along the entire presumptive wing margin, we do not observe the expected transformation from anterior to posterior fate (see Fig. 3c, f). It is possible that high levels of En in the margin are not permissive for posterior fate specification. *c*, UAS-En misexpressed using the 30A-GAL4 driver activates ectopic *dpp* in nearby. Because of the folding of the disc epithelium in this region, the patterns appear overlapping. The 30A-GAL4 driver is expressed in a ring of cells in the wing hinge which encircles the wing pouch (see refs 7, 8, 29). *d-f*, Double labelling to visualize *en* protein (red) and *dpp-lacZ* expression (green). *d* Wild type. *e*, UAS-En misexpressed using the Vg-GAL4 driver. *dpp-lacZ* is expressed outside the regions where En is detected in this disc. Note the ectopic expression of *dpp* in the patch of cells lacking En/Inv in the posterior compartment (arrow). *f*, UAS-En misexpressed using the 30A-GAL4 driver. *g-i*, Double labelling to visualize *en* protein (red) and *ci* protein (green). *g*, The patterns of En and Ci are complementary in the wild type. *h*, UAS-En misexpressed using the Vg-GAL4 driver. Ci is repressed in anterior cells



that express En. Note that the low levels of ectopic En flanking the D-V boundary can repress Ci expression, but comparable levels on *inv* protein cannot (see Fig. 2h for comparison). *i*, UAS-En misexpressed using the 30A-GAL4 driver represses Ci.

reversal of symmetry along the normal vein 2 (Fig. 3c, d). These phenotypes resemble the effects of ectopic expression of *hh* or *dpp* in the anterior compartment^{4,7,8,20}. Unlike ectopic *Inv*, ectopic expression of En causes ectopic expression of *hh* and *dpp* (Fig. 4a-f) and repression of *ci* (Fig. 4g-i); *hh* is misexpressed in a broad stripe overlapping along the D-V boundary (Fig. 4b) and induces *dpp* expression in nearby cells (Fig. 4e).

In more severely affected wings we observe anterior to posterior transformation of the wing hinge (Fig. 3f). Although the transformation to posterior cell identity is complete in the hinge, the wing margin retains a predominantly anterior identity (Fig. 3f). Thus, although En is expressed at quite high levels along the margin (see Fig. 4b, e, h), cell fates are not affected as much as they are by the relatively low level of ectopic *inv* protein produced by the *inv*^D mutation. Note that the *vg-Gal4* driver can also direct low levels of *en* protein in a broad region of the wing blade in some discs (Fig. 4b). This low level of *en* protein is sufficient to activate *hh* expression (Fig. 4b) but, based on the range of adult phenotypes observed, it apparently does not cause transformation to posterior identity in the adult wing blade. Although a direct comparison between the level of *en* protein in such discs and the level of *inv* protein produced by the *inv*^D mutation is difficult, these observations suggest that *en* and *inv* proteins differ in their ability to regulate distinct sets of target genes.

A striking feature of En overexpression is an apparently paradoxical transformation of posterior to anterior in the wing margin (Fig. 3e, f) and in the wing blade (Fig. 3g). These phenotypes may be explained by the observation that high levels of En can cause repression of the endogenous *en* and *inv* genes in the poste-

rior compartment (Fig. 4b, e, h). The absence of the endogenous (and in many cases exogenous) gene product allows posterior cells to adopt anterior identity, as indicated by ectopic expression of *dpp* and *ci* in the posterior compartment (arrows in Fig. 4e, h). Recent evidence suggests that *polyhomeotic*, a member of the Polycomb group of genes, may be a target for transcriptional activation by En²¹. The Polycomb group of genes are responsible for repression of *en* and *inv* expression in the anterior compartment²². It is possible that overexpression of En interferes with maintenance of *en* and *inv* expression through overexpression of *polyhomeotic* in the posterior compartment.

Genetic evidence suggests that the functions of *en* and *inv* are partly overlapping, in that *en* can compensate for the loss of *inv* function in an *inv* null mutant¹⁸. Furthermore, complete transformation of posterior to anterior fate can only result when both *en* and *inv* are removed¹⁶⁻¹⁸. Removing *en* function alone causes only partial transformation of morphological fate from posterior to anterior^{14,15}; *inv* must therefore contribute to posterior fate specification. Consistent with these observations we have shown that *inv* is sufficient to transform anterior to posterior cell fate, without affecting the expression of effector genes, such as *hh*, through which *en* functions to establish the A-P organizer. Conversely, ectopic expression of En leads to misexpression of *hh* and *dpp* without always causing a clear transformation of morphological identity. These findings suggest that the specification of posterior cell identity can be distinguished genetically from the specification of the compartment boundary as a pattern organizing centre. En and *Inv* appear to have different potency in regulating the specific target genes that mediate these distinct functions. The functions of the two *en*-like genes

in vertebrates appear to be more divergent, although analysis of double mutants suggests that they retain partly overlapping functions²³. □

Received 24 May; accepted 16 June 1995.

1. García-Bellido, A., Ripoll, P. & Morata, G. *Nature new Biol.* **245**, 251–253 (1973).
2. Lawrence, P. A. & Morata, G. *Cell* **78**, 181–189 (1994).
3. Diaz-Benjumea, F. J. & Cohen, S. M. *Cell* **75**, 741–752 (1993).
4. Basler, K. & Struhl, G. *Nature* **368**, 208–214 (1994).
5. Tabata, T. & Kornberg, T. *Cell* **76**, 89–102 (1994).
6. Diaz-Benjumea, F. J., Cohen, B. & Cohen, S. M. *Nature* **372**, 175–179 (1994).
7. Capdevila, J. & Guerrero, I. *EMBO J.* **13**, 4459–4468 (1994).
8. Ingham, P. W. & Fietz, M. J. *Curr. Biol.* **5**, 432–440 (1995).
9. García-Bellido, A. & Santamaria, P. *Genetics* **72**, 87–104 (1972).
10. Morata, G. & Lawrence, P. A. *Nature* **255**, 614–617 (1975).
11. Lawrence, P. A. & Morata, G. *Dev. Biol.* **50**, 321–337 (1976).
12. Kornberg, T., Siden, I., O'Farrell, P. & Simon, M. *Cell* **40**, 45–53 (1985).
13. Williams, J. A., Paddock, S. W., Vorwerk, K. & Carroll, S. B. *Nature* **368**, 299–305 (1994).
14. Lawrence, P. A. & Struhl, G. *EMBO J.* **7**, 827–833 (1982).
15. Gubb, D. *Wilhelm Roux Arch. dev. Biol.* **194**, 236–246 (1985).
16. Hidalgo, A. *Curr. Biol.* **4**, 1087–1098 (1994).
17. Sanicola, M., Sekelsky, J., Elson, S. & Gelbart, W. M. *Genetics* **139**, 745–756 (1995).

18. Tabata, T., Schwartz, C., Gustavson, E., Ali, Z. & Kornberg, T. B. *Development* (in the press).
19. Eker, R. *Hereditas* **12**, 217–222 (1929).
20. Lepage, T., Cohen, S. M., Diaz-Benjumea, F. J. & Parkhurst, S. M. *Nature* **373**, 711–715 (1995).
21. Serrano, N. et al. *Development* **121**, 1691–1703 (1995).
22. Busturia, A. & Morata, G. *Development* **104**, 713–720 (1988).
23. Wurst, W., Auerbach, A. B. & Joyner, A. L. *Development* **120**, 2065–2075 (1994).
24. Patel, N. H. et al. *Cell* **58**, 955–968 (1989).
25. Williams, J. A., Bell, J. B. & Carroll, S. B. *Genes Dev.* **5**, 2481–2495 (1991).
26. Schwartz, C., Locke, J., Nishida, C. & Kornberg, T. B. *Development* **121**, 1625–1635 (1995).
27. Blackman, R. K., Sanicola, M., Raftery, L. A., Gillevet, T. & Gelbart, W. M. *Development* **111**, 657–665 (1991).
28. Lee, J. J., von Kessler, D. P., Parks, S. & Beachy, P. A. *Cell* **71**, 33–50 (1992).
29. Brand, A. & Perrimon, N. *Development* **118**, 401–415 (1993).
30. Poole, S. J., Kauvar, L. M., Drees, B. & Kornberg, T. *Cell* **40**, 37–43 (1985).
31. Williams, J. A. et al. *Genetics* **125**, 833–844 (1990).

ACKNOWLEDGEMENTS. A.J.S. and W.J.B. made equal contributions to this work. We thank S. Scianimanico and R. Bhatnagar for providing technical assistance; T. Kornberg for sharing data on the *invected* null mutant phenotype; G. Morata for discussion of unpublished data; C. Schwartz for antibody to Ci protein; M. Hoffman for the Vg-GAL 4 driver; and S. Eaton for discussion and for comments on the manuscript. A.J.S. is supported by a predoctoral fellowship from the Alberta Heritage Foundation for Medical Research. W.J.B. is supported by an EMBO long-term fellowship. Work in J.B.'s laboratory is supported by an operating grant from NSERC (Canada).

Specification of neuroblast identity in the *Drosophila* embryonic central nervous system by *gooseberry-distal*

James B. Skeath, Yu Zhang*, Robert Holmgren*, Sean B. Carroll† & Chris Q. Doe

Howard Hughes Medical Institute and Department of Cell and Structural Biology, University of Illinois, Urbana, Illinois, 61801, USA

* Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208, USA

† Howard Hughes Medical Institute, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

THE *Drosophila* central nervous system develops from a segmentally reiterated array of 30 unique neural precursors, called neuroblasts. Each neuroblast goes through a stereotyped cell lineage to produce an invariant clone of neural progeny. It is critical to identify the genes that specify neuroblast identity as these genes control the time of formation, gene expression profile, and cell lineage characteristics of each neuroblast. Here we show that the *Pax-type* *gooseberry-distal* gene specifies row 5 neuroblast identity. Initially, four rows of neuroblasts form per segment (1, 3, 5, 7) and *gooseberry-distal* is expressed in row 5 neuroblasts^{1–3}. By using 10 molecular markers, and by following the number and orientation of neuroblast divisions, we show that lack of *gooseberry-distal* transforms row 5 neuroblasts into row 3 neuroblasts, whereas ubiquitous *gooseberry-distal* generates the reciprocal transformation. Thus, *gooseberry-distal* is necessary and sufficient to specify row 5 neuroblast identity autonomously. The 10 genes coordinately regulated by *gooseberry-distal* are prime candidates for controlling specific aspects of neuroblast identity.

In the *Drosophila* central nervous system (CNS), neuroblasts delaminate in a fixed spatiotemporal sequence from a ventral neuroectoderm to form a stereotyped sub-epidermal array. The first 10 (S1) neuroblasts segregate about 30 min after gastrulation in an orthogonal array of four rows (1, 3, 5, 7) and three columns (medial, intermediate, lateral) per hemisegment. The even-numbered neuroblast rows form slightly later in development. After delaminating, each neuroblast follows a unique cell lineage to produce on average five ganglion mother cells (GMCs); each GMC divides yielding a characteristic pair of neurons or glia⁴. Neuroblast formation depends on the function of the *achaete-scute* complex of proneural genes: *achaete* (*ac*),

scute (*sc*) and *lethal of scute* (*lsc*)⁵. These genes are initially expressed within the neuroectoderm in partially overlapping arrays of 'proneural clusters' of about 6 cells^{6–8}. One cell per cluster retains proneural gene expression, delaminates and becomes the neuroblast. Although the genetic regulatory mechanisms that govern neuroblast formation are well understood, those that control neuroblast identity are poorly understood.

The segment polarity genes are excellent candidates to govern neuroblast identity: most are regionally expressed in the neuroectoderm at the time of neuroblast formation⁹ and mutations in them often perturb CNS development^{3,10,11}. The *gooseberry* (*gsb*) locus consists of two tightly linked genes: *gsb-distal* (*gsb-d*) and *gsb-proximal* (*gsb-p*), which encode transcription factors with homeobox and paired box DNA-motifs^{12,13}. Both genes are structurally similar to vertebrate *Pax* genes, which are regionally expressed in the CNS and required for proper CNS development¹⁴. *gsb-d* transcription initiates in row 5 and 7 neuroectodermal cells before neuroblasts form and is rapidly restricted to row 5 neuroectoderm, row 5 neuroblasts and neuroblast 7-1 (refs 2, 3). *gsb-p* is activated well after S1 neuroblasts form (late stage 10) in the same neuroblasts as *gsb-d*². To determine the role of *gsb* in specifying neuroblast identity, we assayed 10 neuroblast markers and neuroblast cell division patterns in embryos that carry a 90-kilobase (kb) deletion which completely removes both *gsb* transcripts¹⁵ (*gsb*[−] embryos), and in embryos that ubiquitously express *gsb-d* under the *hsp70* promoter (*HS-gsb*). Assays of 8 markers (all but 22C10 and *wingless*) were done in embryos before expression of *gsb-p*, and thus these phenotypes are specifically due to the genetic loss of *gsb-d*.

The first CNS phenotype in *gsb*[−] embryos occurs before any neuroblasts form, and reflects a change from row 5 to row 3 neuroectodermal fate. In wild-type embryos, *ac* and *sc* are coexpressed in two clusters (medial and lateral) in rows 3 and 7 (Fig. 1a; not shown for *sc*). In contrast, *lsc* is not expressed in row 3; it is transcribed in rows 1, 5 and 7 (Fig. 1d)^{6,7}. In *gsb*[−] embryos, the *ac/sc* and *lsc* proneural clusters arise normally^{7,16}, but the patterns change before neuroblasts form. In row 5, the normal *gsb-d* expression domain, *ac/sc* expression turns on and *lsc* turns off (Fig. 1b, e); this reflects a transformation from a row 5 to a row 3 pattern of proneural gene expression. The reciprocal neuroectodermal phenotype (row 3 is transformed to row 5) is observed in *HS-gsb* embryos: row 5 *lsc* expression expands anteriorly, replacing that of *ac/sc* in row 3 (Fig. 1c, f). In wild-type, *gsb*[−], and *HS-gsb* embryos, each S1 neuroblast retains the proneural gene expression profile of the cluster from which it segregates (Fig. 1g–k). Thus, the alterations observed to the neuroectodermal expression patterns of the proneural

genes in *gsb*⁻ and HS-*gsb* embryos are conserved in S1 neuroblasts. Both *gsb*⁻ and HS-*gsb* embryos have subtle changes to proneural gene expression occurring in row 7 neuroectoderm and neuroblasts (Fig. 1), presumably because of transient *gsb-d* expression in row 7.

The changes to proneural gene expression in *gsb*⁻ and HS-*gsb* embryos are the earliest indicators of altered neuroectoderm and neuroblast identity; to confirm and extend these results, we examined additional neuroblast markers and assayed the division profile of selected neuroblasts. For simplicity, we focus on the role of *gsb-d* in specifying differences between the medial row 3 'neuroblast', called MP2, and the medial row 5 neuroblast, called 5-2. These two cells can be distinguished by gene expression and by the number and orientation of their cell divisions.

In wild-type embryos, MP2 expresses *fushi-tarazu* (*ftz*; Fig. 2a), *odd-skipped* (*odd*; Fig. 2b), nuclear prospero protein, the AJ96 enhancer trap, and the 22C10 epitope (not shown); in addition, MP2 is the only 'neuroblast' that divides once, symmetrically, to produce two neurons (inset, Fig. 2b)¹⁷. In contrast, neuroblast 5-2 expresses *seven-up-lacZ* (*suv-lacZ*; Fig. 2c), *wingless* (*wg*; not shown), and repeatedly divides asymmetrically to generate a series of GMCs (inset, Fig. 2c).

In *gsb*⁻ embryos, neuroblast 5-2 forms 56% of the time; lowered frequency of formation is probably due to altered proneural gene expression (see above). It expresses all seven row 3 markers but no row 5 markers, and divides once symmetrically like MP2 (Fig. 2d-f). The coordinate switch from row 5 to row 3 markers, plus the MP2-like symmetrical cell division, is strong evidence

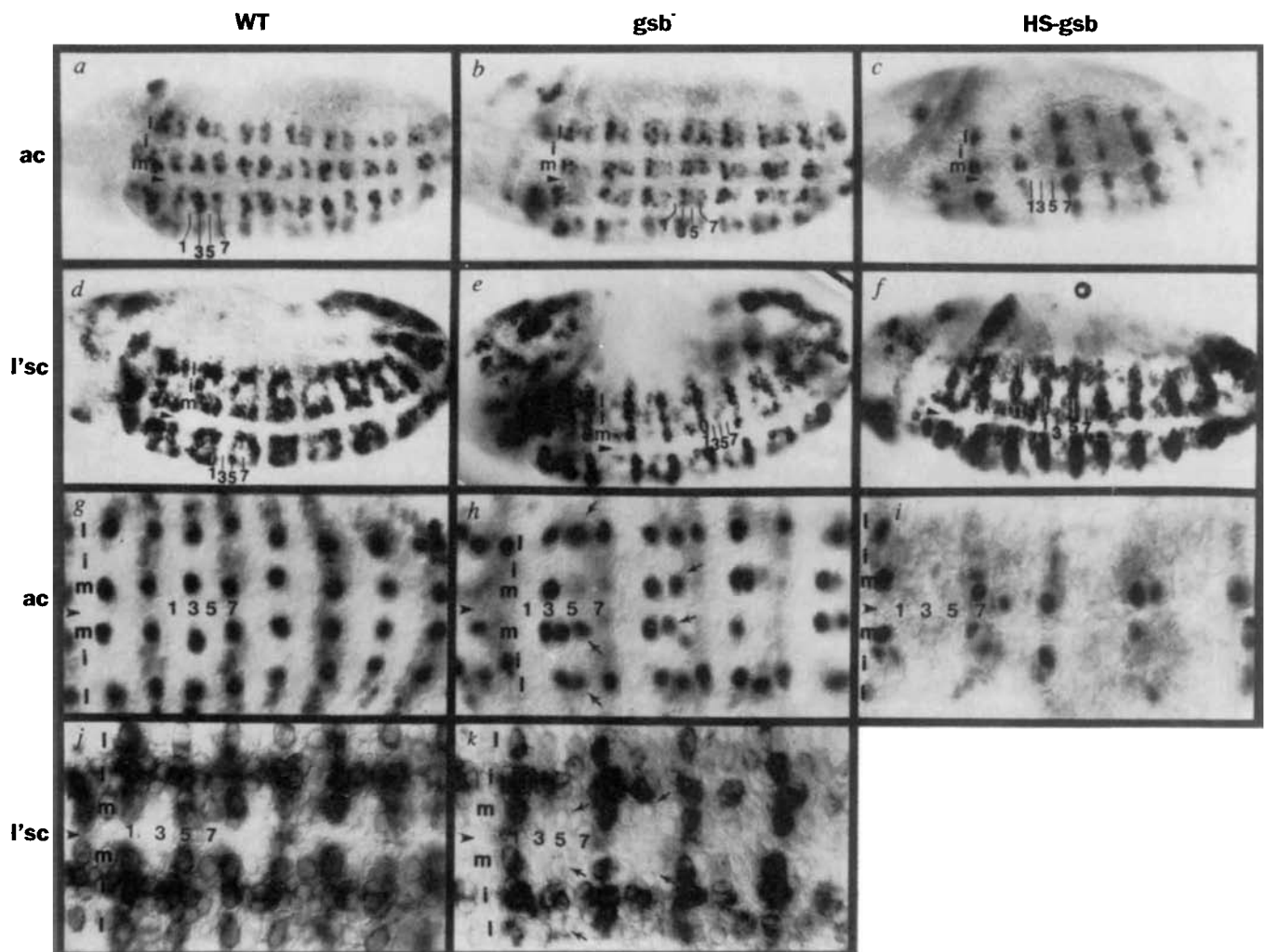


FIG. 1 *gsb-d* governs proneural gene expression in neuroectodermal clusters and S1 neuroblasts. *ac* protein and *l'sc* RNA expression in wild-type (a, d, g, j), *gsb*⁻ (b, e, h, k) and HS-*gsb* (c, f, i) embryos in the neuroectoderm (a-f; stage 8) and in S1 neuroblasts (g-k). In all panels, rows 1, 3, 5 and 7 and the medial (m), intermediate (i) and lateral (l) neuroblast columns are indicated; anterior, left; arrowhead, ventral midline; in panels g-i, *engrailed* (*en*) expression marks row 7. In wild-type embryos (WT), *ac* is expressed in row 3 and 7 proneural clusters (a); *l'sc* in those of rows 1, 5 and 7 (d). The S1 neuroblasts that form from these clusters retain this profile of proneural gene expression: row 3 and 7 neuroblasts express *ac* (g); row 1, 5 and 7 neuroblasts express *l'sc* (j). Once S1 neuroblasts form, neuroectodermal cells in the intermediate column activate *l'sc* transcription (j). S2 neuroblasts arise from these *l'sc* positive clusters. In *gsb*⁻ embryos, row 5 neuroectodermal cells activate *ac* (b) but extinguish *l'sc* gene expression (e); subsequently

row 5 neuroblasts express *ac* (arrows, h) instead of *l'sc* (arrows, k). Furthermore, row 7 neuroectodermal cells show a decrease in *ac* and *l'sc* gene expression (b, e) and row 7 neuroblasts prematurely extinguish *ac* and *l'sc* gene expression (h, k). Again, as a prelude to S2 neuroblast segregation, cells in the intermediate column activate *l'sc* transcription (k). In HS-*gsb* embryos, row 3 neuroectodermal cells turn off *ac* (c) and turn on *l'sc* gene expression (f), this is most pronounced in the m and l columns. Subsequently, row 3 neuroblasts form but do not express *ac*, and duplicated *ac*-positive neuroblasts occasionally arise in row 7 (i). As S1 neuroblasts segregate, ubiquitous *gsb* expression induces a nearly uniform expression of *l'sc* within the neuroectoderm, presumably reflecting *l'sc*'s role in S2 neuroblast formation, this precludes the identification of *l'sc*-positive S1 neuroblasts.

METHODS. Antibody and RNA *in situ* methods are described in ref. 16. For heat-shock regimens, see Fig. 2.

that *gsb-d* is required to specify 5-2 identity (instead of a default MP2 identity). The opposite transformation occurs in HS-*gsb* embryos: 'MP2' expresses all row 5 markers; none of the row 3 markers tested, and divides asymmetrically like neuroblast 5-2 (Fig. 2g-i). These data show that *gsb* is sufficient to induce ectopic neuroblast 5-2 identity at the expense of the MP2 cell fate. Thus, molecular markers and cell division patterns reveal a consistent 5-2 to MP2 (row 5 to 3) transformation in *gsb*⁻ embryos and a MP2 to 5-2 (row 3 to 5) transformation in HS-*gsb* embryos.

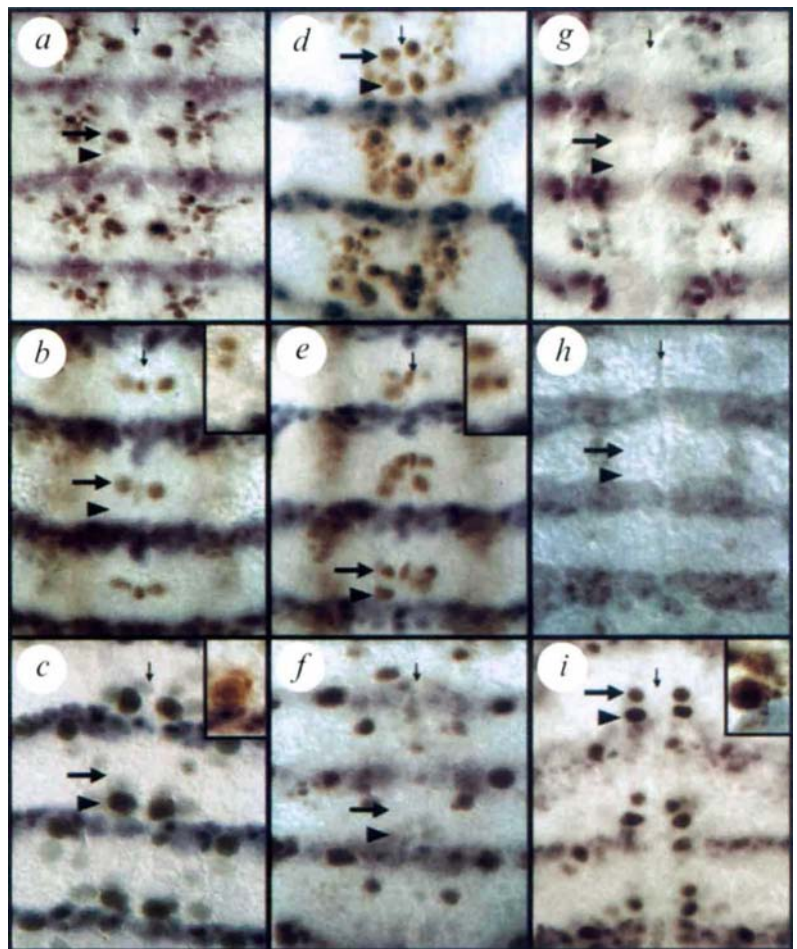
We chose to analyse MP2 and 5-2 in detail because of the abundance of available markers for these neuroblasts and the differences in their division patterns. However, available markers for intermediate and lateral neuroblasts (*ac*, *sc*, *lsc*, *wg* and *svp-lacZ*) show identical transformations in *gsb*⁻ and HS-*gsb* embryos. Furthermore, the intermediate column row 3 neuroblast, called neuroblast 4-2, forms just after the S1 neuroblasts and produces the identified motor neuron RP2¹. In *gsb*⁻ embryos there are two RP2 neurons¹⁰, consistent with a row 5 to 3 transformation; conversely, in HS-*gsb* embryos no RP2 neurons form³, consistent with a row 3 to 5 transformation. This is a second example of a *gsb*-dependent transformation of neuroblast fate assayed by cell division profile or cell lineage (in addition

to molecular markers). Taken together with our results, these data strongly imply that *gsb-d* specifies the identity of row 5 neuroblasts and represses the identity of row 3 neuroblasts (Fig. 3).

The earliest effect of *gsb-d* is on neuroectodermal expression of proneural genes, thus we speculate that *gsb-d* governs neuroblast identity by specifying cell fate in the neuroectoderm. To test directly when *gsb-d* specifies MP2 versus 5-2 identity, we ectopically expressed *gsb* either before neuroblast formation, or just after formation but before division, and assayed for development of a normal *ftz*-positive MP2 (Fig. 2, legend). We found that ectopic *gsb* transforms MP2 into 5-2 when expressed in the neuroectoderm before neuroblast formation: only 7% of the hemisegments contain the normal *ftz*-positive MP2 ($n=108$). In contrast, ectopic *gsb* does not alter MP2 fate when ectopically expressed after MP2 formation but before its division: 85% of the hemisegments contain a *ftz*-positive MP2 ($n=110$). These results, taken together with the observation that *gsb-d* regulates proneural gene expression in the neuroectoderm, show that *gsb-d* controls neuroblast identity by specifying cell fate in the neuroectoderm; that is, neuroblast identity is assigned before the 'singling out' of a neuroblast from the neuroectoderm. In addition, these results raise the possibility that the proneural genes may

FIG. 2 *gsb-d* is necessary and sufficient to specify row 5 neuroblast identity. *ftz* (a, d, g), *odd* (b, e, h) and *svp-lacZ* (c, f, i) gene expression in wild-type (a, b, c), *gsb*⁻ (d, e, f) and HS-*gsb* (g, h, i) embryos. *en* is used as a positional marker for row 7; anterior, up; midline, small arrow. In wild-type embryos, the row 3 MP2 (arrow) expresses *ftz* and *odd* (a, b) and divides nearly symmetrically (inset, b); the row 5 neuroblast 5-2 (arrowhead) expresses *svp-lacZ* (c) and divides asymmetrically (inset, c) multiple times. In *gsb*⁻ embryos, both row 3 (arrow) and row 5 (arrowhead) medial neuroblasts express the row 3 markers *ftz* and *odd* (d, e), do not express the row 5 marker *svp-lacZ* (f), and divide nearly symmetrically (shown for the row 5 'MP2'; inset, e). In HS-*gsb* embryos, both row 3 (arrow) and row 5 (arrowhead) medial neuroblasts lack the row 3 markers *ftz* and *odd* (g, h), express the row 5 marker *svp-lacZ* (i), and divide asymmetrically (inset, i). The *ftz* and *odd* stained embryos are stage 10; the *svp-lacZ* stained embryos are stage 9.

METHODS. Neuroblast formation in *gsb*⁻ embryos: 5-2, 56% (62/111 hemisegments), 5-3, 53% (19/30), 5-6, 48% (11/23). Gene expression in wild-type embryos: MP2, *ftz*-positive >99% (111/112); 5-2, *svp-lacZ*-positive 100% (134/134); 5-3 *svp-lacZ*-positive 98% (83/85); 5-6, *svp-lacZ*-positive 89% (64/72); gene expression in *gsb*⁻ embryos: 5-2, *ftz*-positive 52% (75/145), *svp-lacZ*-positive <1% (3/368); 5-3 and 5-6, *svp-lacZ*-positive 21% (25/119) and 19% (14/76), respectively. Gene expression in HS-*gsb* neuroblasts: MP2, *svp-lacZ*-positive 50% (31/64) and *ftz*-positive 7% (8/108). Furthermore, in *gsb*⁻ embryos all row 5 neuroblasts extinguish *wg* expression by stage 11; whereas in HS-*gsb* embryos all row 3 and 5 neuroblasts express *wg* (not shown; ref. 3). Antibody double labelling was performed as in ref. 16. Heat-shock experiments: embryos were collected for 30 min, grown for various times at 23 °C (see below), heat shocked at 38 °C for 30 min, developed at 23 °C for 1-3 h, then fixed and assayed for normal (*ftz*-positive) MP2 development. A subset of embryos were stained for *ac* before heat shock to determine the stage of the embryos (that is, whether MP2 had formed). For 3-3.5 h heat-shocked embryos: before heat shock MP2 has not formed in >90% of the hemisegments; after heat-shock treatment only 8/108 (7%) hemisegments have a *ftz*-positive MP2. For 3.25-3.75 h heat-shocked embryos: before heat shock MP2 has not formed in ~90% of the hemisegments; after heat shock only 24/112 (21%) hemisegments contain a *ftz*-positive MP2. For 3.75-4.25 h heat-shocked embryos: before heat shock MP2 has formed in ~50% of the hemisegments; after heat shock a *ftz*-positive MP2 was found in 106/278 (38%) hemisegments. For 4-4.5 h



heat-shocked embryos: before heat shock MP2 had formed in ~90% of the hemisegments; after heat shock 93/112 (85%) hemisegments contained a *ftz*-positive MP2. Wild-type embryos that underwent identical heat-shock regimens and untreated HS-*gsb* embryos showed no loss of *ftz* or duplication of *svp-lacZ* expression in MP2 or 5-2 (>98% of hemisegments; $n>100$ for each).

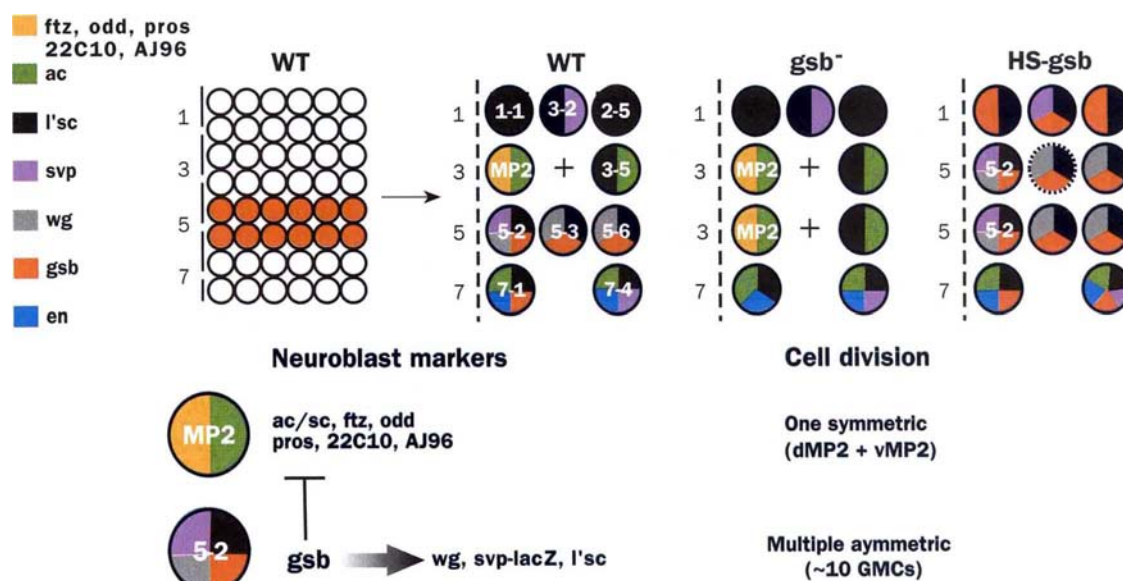


FIG. 3 Top, *gsb-d* is necessary and sufficient for row 5 neuroblast specification. Far left, molecular markers used to assay neuroblast identity in wild-type, *gsb*⁻ and HS-*gsb* embryos. All markers were assayed in wild-type and *gsb*⁻ embryos and all markers except AJ96, *pros* and 22C10 were assayed in HS-*gsb* embryos. +, Presence of the intermediate column row 3 S2 neuroblast, neuroblast 4-2, which forms just after S1 neuroblasts and produces the identified motor neuron RP2. In wild-type embryos, *gsb-d* is expressed in row 5 neuroectoderm and neuroblasts. In *gsb*⁻ embryos, row 3 neuroblasts are transformed into row 5 neuroblasts. In HS-*gsb* embryos, ubiquitous *gsb* causes the reciprocal transformation. Stippled circle around neuroblast indicates that the

neuroblast forms at this location <50% of the time. Anterior, top; ventral midline, dashed line. Neuroblast nomenclature is revised from ref. 1 according to J. Broadus, J.B.S. and C.Q.D. (unpublished observations). Bottom, *gsb-d* is at the top of a gene regulatory hierarchy that specifies the fate of neuroblast 5-2. In neuroblast 5-2 *gsb-d* regulates positively the *wg*, *l'sc* and *svp-lacZ* genes but represses the *ac/sc*, *ftz*, *odd*, *pros*, 'AJ96' and '22C10' genes, which are expressed in MP2. These two sets of genes probably act coordinately to govern the fate of neuroblast 5-2, which divides repeatedly to yield ~10 GMCs, and MP2, which divides once to produce two neurons: dMP2 and vMP2.

help confer neuroblast identity, because they are expressed in the neuroectoderm at the time when neuroblast identity is specified.

Absence of *gsb-d* transforms row 5 to 3, whereas ectopic *gsb-d* transforms row 3 to 5. Yet, ectopic *gsb-d* has relatively little effect on row 1 and 7 neuroblasts (Fig. 3), suggesting that only rows 3 and 5 are competent to respond to *gsb-d*. Each half-segment (rows 3/5 versus rows 1/7) is distinguished by pair-rule gene expression: rows 3 and 5 express low levels of *ftz* or *even-skipped* (*eve*), and comprise the 'wg-competent' domain; rows 1 and 7 express high levels of *ftz* or *eve*, and comprise the 'engrailed-competent' domain¹⁸. Low levels of *ftz* and *eve* may provide a 'ground state' row 3 fate, with *gsb-d* expression able to induce row 5 identity within this domain.

gsb-d is necessary and sufficient to specify neuroblast 5-2 identity and, most probably, the identity of all other row 5 neuroblasts (Fig. 3). *gsb-d* appears to be at the top of a gene regulatory

hierarchy: it coordinately activates at least three genes in neuroblast 5-2 and coordinately represses at least seven genes normally expressed in MP2. We envisage three ways *gsb-d* might confer row 5 neuroblast identity. First, *gsb-d* regulates proneural gene expression, and proneural genes control both neuroblast formation and specification. Second, *gsb-d* independently regulates proneural genes to control neuroblast formation, and candidate neuroblast identity genes (like *svp* and *odd*) to control neuroblast identity. Third, *gsb-d* independently regulates proneural and candidate neuroblast identity genes; both classes of genes then act together to govern neuroblast identity. Preliminary results show that loss of *ac/sc* function partially transforms the symmetrically dividing MP2 into an asymmetrically dividing neuroblast (J.B.S. and C.Q.D., unpublished observations); this supports models where proneural genes contribute to the specification of neuroblast identity. □

Received 12 May; accepted 23 June 1995.

1. Doe, C. Q. *Development* **119**, 855–863 (1992).
2. Gutjahr, T., Patel, N. H., Li, X., Goodman, C. S. & Noll, M. *Development* **119**, 21–31 (1993).
3. Zhang, Y., Ungar, A., Fresquez, C. & Holmgren, R. *Development* **120**, 1151–1161 (1994).
4. Goodman, C. S. & Doe, C. Q. in *The Development of Drosophila* (eds Bate, M. & Martinez-Arias, A.) 1131–1206 (Cold Spring Harbor Laboratory Press, New York, 1993).
5. Jimenez, F. & Campos-Ortega, J. A. *Neuron* **5**, 81–89 (1990).
6. Cabrera, C. V., Martinez-Arias, A. & Bate, M. *Cell* **50**, 425–433 (1987).
7. Martin-Bermudo, M. D., Gonzalez, F., Dominguez, M., Rodriguez, L. & Jimenez, F. *Development* **119**, 1003–1012 (1991).
8. Skeath, J. B. & Carroll, S. B. *Development* **114**, 939–946 (1992).
9. Akam, M. *Development* **101**, 1–21 (1987).
10. Patel, N. H., Schafer, B., Goodman, C. S. & Holmgren, R. *Genes Dev.* **3**, 890–904 (1989).
11. Chu-LaGriff, Q. & Doe, C. Q. *Science* **261**, 1594–1597 (1993).

12. Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. & Noll, M. *Cell* **47**, 1033–1040 (1986).
13. Baumgartner, S., Bopp, D., Burri, M. & Noll, M. *Genes Dev.* **1**, 1247–1267 (1987).
14. Chalepakidis, G., Stoykova, A., Wijnholds, J., Tremblay, P. & Gruss, P. *J. Neurobiol.* **24**, 1367–1384 (1993).
15. Cote, S. et al. *EMBO J.* **6**, 2793–2801 (1987).
16. Skeath, J. B., Panganiban, G. F. & Carroll, S. B. *Genes Dev.* **6**, 2606–2619 (1992).
17. Doe, C. Q., Hiromi, Y., Gehring, W. J. & Goodman, C. S. *Science* **239**, 190–195 (1988).
18. Ingham, P. W. & Martinez-Arias, A. *Cell* **68**, 221–235 (1992).

ACKNOWLEDGEMENTS. We thank N. Patel for comments on the manuscript; E. Ward and D. Coulter for generously providing us with *odd-skipped* antibody before publication. This work was supported by a Damon Runyon-Walter Winchell Cancer Research Fellowship to J.B.S., an NSF PYI Award to S.B.C.; an NIH grant to R.H.; an NIH grant and NSF PYI award to C.Q.D. C.Q.D. and S.B.C. are assistant and associate investigators of the Howard Hughes Medical Institute, respectively.

Resistance to herpes stromal keratitis conferred by an IgG2a-derived peptide

Anne C. Avery*, Zi-Shan Zhao*,
Alejandro Rodríguez†, Elizabeth K. Bikoff‡,
Masoud Soheillan†, C. Stephen Foster†
& Harvey Cantor*

* Department of Pathology, Harvard Medical School, Laboratory of Immunopathology, Dana-Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts 02115, USA

† Hiles Immunology Laboratory, Massachusetts Eye and Ear Infirmary, Department of Ophthalmology, Harvard Medical School, 243 Charles Street, Boston, Massachusetts 02114, USA

‡ Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

NOT all peripheral tissue antigens enter the thymus and it is unclear how the immune system remains tolerant to this class of self antigen. As tolerance to self peptides can generate gaps in the T-cell repertoire for cross-reactive foreign antigens^{1,2}, we investigated whether this mechanism might also diminish autoimmune reactions to similar peptides expressed by peripheral tissues. Herpes stromal keratitis (HSK) is a virally induced autoimmune reaction against corneal tissues mediated by T cells³⁻⁵, and is a leading cause of human blindness⁶. Resistance to HSK in mice is associated with allotypic variation in immunoglobulin genes^{7,8}, possibly because circulating immunoglobulin-derived peptides can cross-tolerize T cells specific for corneal tissue autoantigens. Here we show that HSK is mediated by T-cell clones specific for corneal self antigens which also recognize an allotype-bearing peptide derived from IgG2a, and that exposure of HSK-susceptible mice to a soluble form of this peptide confers resistance to HSK. Shared expression of peptide subsequences between sequestered tissue proteins and circulating proteins may be important for maintenance of self-tolerance and prevention of autoimmunity.

Inbred mouse strains that carry the *Igh*^{d/c} alleles (for example, C.AL-20 (H-2^d, *Igh*^d)) are susceptible to necrotizing stromal keratitis induced by infection with herpes simplex virus 1 (HSV-1) (strain KOS), whereas congenic strains that carry the *Igh*^b allele (for example, C.B-17 (H-2^d, *Igh*^b)) are highly resistant⁹. We investigated whether exposure to antigens derived from *Igh*^b (but not *Igh*^{d/c}) immunoglobulins resulted in *Igh*^b-linked resistance to HSK. We injected immunoglobulin, prepared to induce tolerance¹⁰, from serum of either C57BL/6 (*Igh*^b) or A/J (*Igh*^d)

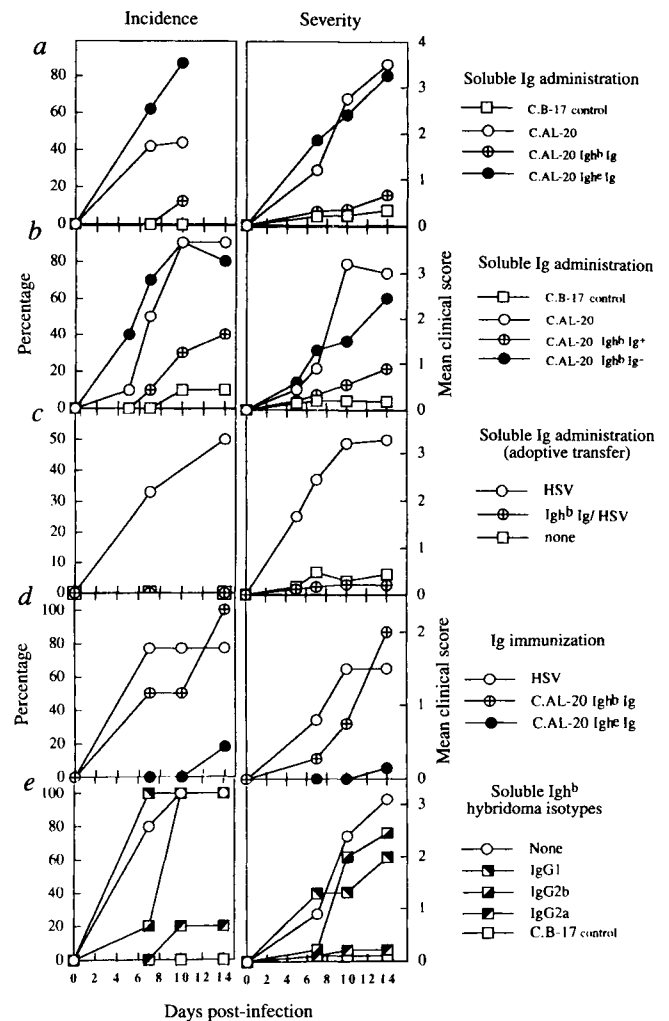
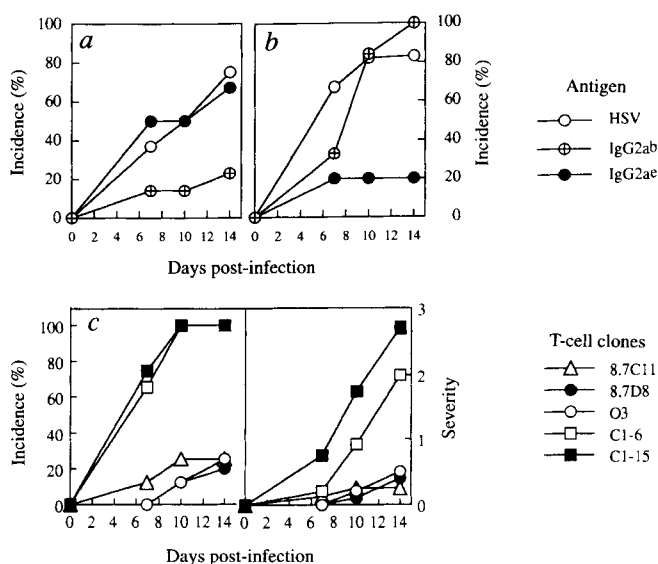


FIG. 1 a, Administration of deaggregated partly purified *Igh*^b but not *Igh*^d prevents HSK. C.AL-20 mice were injected intravenously with 1 mg immunoglobulin (Ig) from C57BL/6 serum (*Igh*^b; crossed circles) or A/J serum (*Igh*^d; filled circles) at days -14 and -7 before corneal HSV-1 infection. Control C.AL-20 (open circles) and C.B-17 (open squares) mice were untreated before HSV-1 infection. Each data point represents 8 mice. Disease incidence (left) and mean clinical score (right) are shown. Partly purified B6 and A/J Ig obtained after precipitation with 45% ammonium sulphate (pH 7.4) was deaggregated by centrifugation (10,000 g for 30 min at 4 °C) before administration according to the schedule described above. Briefly, after anaesthesia the right eye of each mouse was scratched with a 23G needle and 5 µl 5×10^4 plaque-forming units (PFU) of HSV-1 (strain KOS) was deposited onto the scarified cornea. The severity of clinical stromal keratitis was scored on a scale of 0-4+ at days 7, 10 and 14 as follows. Corneal opacity and neovascularization over less than 25% of the cornea, 1+; less than 50%, 2+; less than 75%, 3+; and 75-100%, 4+ (ref. 25). b, Ig-depleted serum fractions do not prevent HSK. *Igh*^b Ig from C57BL/6 serum was obtained by sequential elution from a protein A column and a rabbit anti-mouse Ig Sepharose 4B column. Protein A-purified Ig (500 µg) (crossed circles) or Ig-depleted serum (500 µg) (filled circles) was given

intravenously 2 weeks and intraperitoneally 1 week before corneal infection with HSV-1. Control C.AL-20 (open circles) and C.B-17 (open squares) mice were untreated before infection. Each point represents 10 mice. Disease incidence (left) and mean clinical score (right) are shown. Contaminating Ig in Ig-depleted serum was less than 0.1 µg per 500 µg, and Ig preparations contained less than 5-10 µg of contaminating serum proteins per 500 µg. c, Intravenous administration of Ig abrogates the ability of T cells to transfer HSK. C.AL-20 mice were given protein A-purified soluble *Igh*^b Ig 2 weeks before corneal infection with HSV-1 (5×10^4 PFU). Two weeks later, 7×10^6 T cells from Ig-treated HSV-1-infected mice (crossed circles), untreated HSV-1-infected mice (open circles) or naive mice (open squares) were injected intraperitoneally into BALB/c nu/nu recipients (6 per group) which were simultaneously infected in the cornea with HSV-1. After passage over Collect columns (Biotex, Edmonton, Canada), T cells were 90-95% CD3⁺, and <1% B220⁺ or class-II MHC⁺. d, T cells from C.AL-20 donors immunized with *Igh*^b protein transfer HSK. C.AL-20 mice were immunized with *Igh*^b (crossed circles) or *Igh*^d (filled circles) Ig (300 µg) in CFA at -3 weeks and in incomplete Freund's at -1 week, or infected with HSV-1 (open circles) before obtaining T cells for adoptive transfer into BALB/c nu/nu recipients immediately before challenge in the cornea with HSV-1. Lymph nodes and spleens from HSV-1-infected (at -2 weeks) mice were used as the source of T cells. e, Effects of different hybridoma proteins on HSK resistance. *Igh*^b hybridoma proteins of the IgG1, IgG2a and IgG2b isotypes were administered intravenously (day -14) and intraperitoneally (day -7) followed by HSV-1 corneal infection. Controls included untreated C.AL-20 (open circles) and C.B-17 (open squares) mice. Cell lines producing these monoclonal antibodies are LS-136 (IgG1^b, specific for mouse lambda light chains), RG9/6.13.HLK (IgG2b^b, specific for rat IgG2a) and RG7/7.6.HL (IgG2a^b, specific for rat kappa light chains). IgG was purified (Mab Trap G, Pharmacia, Piscataway, NJ) and ultracentrifuged before 50 µg IgG was given per injection before intraocular HSV-1 infection.



mice into susceptible C.AL-20 (Igh^d) mice 2 weeks before corneal challenge with HSV-1. Mice of the strain C.AL-20 that received Igh^b- but not Igh^e-containing material did not develop HSV-1-keratitis, and were similar to C.B-17 genetic controls (Fig. 1a). Moreover, intravenous administration of immunoglobulin-depleted serum from C57BL/6 mice into susceptible C.AL-20 mice had no effect on susceptibility to HSK, but protein A-purified immunoglobulin from the same material resulted in almost complete inhibition of HSK (Fig. 1b). This was not due to innate anti-HSK activity, as neither immunoglobulin preparation contained detectable viral neutralizing activity, and only Igh^e immunoglobulin showed a detectable HSV-1 titre in a more sensitive enzyme-linked immunosorbent assay (results not shown).

To determine whether inhibition of HSK in C.AL-20 mice by Igh^b immunoglobulin reflected an effect on the host T-cell response, we examined the ability of T cells from immunoglobulin-treated donors to transfer HSK. BALB/c nu/nu mice, which lack normal T cells, do not develop HSK after corneal HSV-1 infection unless it is reconstituted with T cells from donors immune to HSV-1 (refs 4, 11, 12). As expected, nu/nu recipients

FIG. 2 a, IgG2a^b but not IgG2a^e induces resistance to HSK. Mice were injected with IgG2a monoclonal proteins representing IgG2a^b (RG9/6.13.HLK) (crossed circles) and IgG2a^e (filled circles) (31-62; provided by A. Marshak-Rothstein; see ref. 26) allotypes as described above. HSK incidence and severity after HSV-1 infection was compared with that of HSV-1-infected C.AL-20 mice (open circles). b, T cells from C.AL-20 mice immunized with monoclonal Igh^b can transfer HSK. C.AL-20 mice were immunized with monoclonal IgG2a^b (crossed circles) or IgG2a^e (filled circles) in Freund's as described in the legend to Fig. 1d. T cells from these and from HSV-1-infected mice (open circles) were transferred into BALB/c nu/nu recipients immediately before intraocular HSV-1 infection. c, Adoptive transfer of HSK by CD4 T-cell clones specific for IgG2a^b. BALB/c nu/nu mice were injected with 10⁶ cells from syngeneic BALB/c CD4⁺ Th₁ clones C1-6 (open squares) and C1-15 (filled squares) (specific for IgG2a^b)^{13,27}, 8.7C11 (open triangles) and 8.7D8 (filled circles) (specific for myelin basic protein (MBP))⁶ and O3 (open circles) (specific for ovalbumin (OVA))¹⁵ immediately before intraocular challenge with HSV-1. HSK was measured at days 7, 10 and 14 in each group (3-5 mice per group). Severity measured as mean clinical score.

of T cells from C.AL-20 mice immune to HSV-1 developed HSK after corneal infection with HSV-1 (Fig. 1c). By contrast, nu/nu recipients of T cells from C.AL-20 mice treated with purified Igh^b immunoglobulin before HSV-1 immunization did not develop HSK (Fig. 1c). Igh^b-dependent inhibition of the adoptive T-cell HSK response did not reflect nonspecific reduction of T-cell reactivity, because the *in vitro* response of T cells from treated and untreated mice to HSV-1 and CD3 ligation were not significantly different (not shown).

The hypothesis that HSK is mediated by T-cell clones that recognize Igh^b-derived peptides suggests that immunization of C.AL-20 (Igh^d) mice with immunoglobulin from Igh^b but not Igh^e donors should elicit T cells capable of transferring HSK. To test this, T cells from C.AL-20 mice immunized with Igh^e or Igh^b immunoglobulin emulsified in complete Freund's adjuvant were transferred into BALB/c nu/nu recipients before corneal challenge with HSV-1. Recipients of T cells from C.AL-20 donors immunized with Igh^b developed HSK that was at least as severe as recipients of T cells immune to HSV-1 (Fig. 1d). Mice that received T cells from C.AL-20 donors immunized with Igh^e immunoglobulin did not develop significant HSK (Fig. 1d).

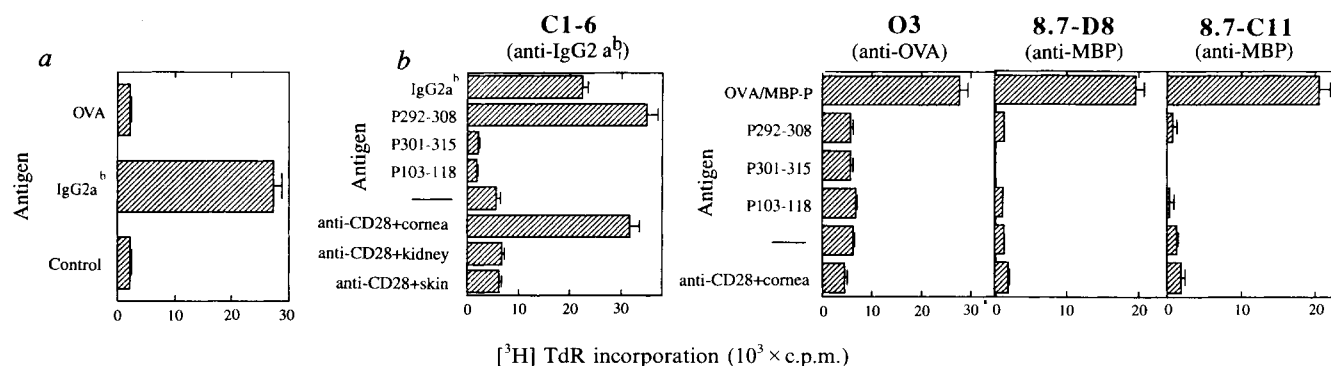


FIG. 3 Specificity profile of IgG2a^b-responsive T-cell clone C1-6. a, Draining periocular lymph nodes were obtained at day 14 after intraocular HSV-1 infection of BALB/c nu/nu recipients that had been reconstituted with clone C1-6 (10⁶ per mouse). Cells from these draining lymph nodes were stimulated with either IgG2a^b (300 µg ml⁻¹) or OVA (300 µg ml⁻¹) for 72 h before [³H]thymidine incorporation was determined. b, Cells from the IgG2a^b-specific T-cell clone C1-6, MBP peptide (61-76)-specific clones 8.7-D8 and 8.7-C11, and OVA-specific clone O3 (2 × 10⁴ cells per well) were incubated with murine corneal protein extracts, murine kidney protein and murine skin protein (500 µg ml⁻¹), or the indicated peptides (0.5 µg ml⁻¹) in the presence of irradiated

(3,300 R) syngeneic BALB/c spleen cells (5 × 10⁵ cells per well) as described¹³. Anti-mouse CD28 (10 µg ml⁻¹) was added to the indicated cultures. Proliferation was assessed after 3 days by a 16-18 h exposure to 1 µCi of [³H] TdR. The results are expressed as mean c.p.m. ± s.e.m. of triplicate cultures. Protein was extracted from murine cornea, kidney and skin by sonication of Tris-buffered cell lysate, as previously²⁸. In addition to the data shown for peptides 301-315 and 103-118 at 0.5 µg ml⁻¹, these peptides did not exhibit detectable stimulatory activity for C1-6 cells at concentrations ranging from 0.1 to 100 µg ml⁻¹. Results essentially identical to clone C1-6 were obtained with C1-15.

The Igh C-region cluster includes genes encoding different heavy-chain isotypes, including IgG1, IgG2a, IgG2b and IgG3. Monoclonal Igh^b hybridoma proteins that expressed different isotypes were tested for their ability to confer resistance to HSK. Intravenous injection of IgG2a^b hybridoma proteins into C.AL-20 mice resulted in inhibition of HSK, but injection of IgG2b or IgG1^b did not (Fig. 1e). In contrast to the effects of IgG2a^b protein, an IgG2a^b hybridoma protein had no inhibitory effect, as expected from the linkage of HSK resistance to the Igh^b allotype (Fig. 2a). We also tested the ability of T cells from mice immunized with IgG2a hybridoma proteins to transfer HSK into nu/nu recipients. BALB/c nu/nu recipients of T cells immune to IgG2a^b, but not IgG2a^e, developed HSK (Fig. 2b).

Previous studies have shown that HSK is mediated by CD4 T cells, as shown by adoptive transfer of HSK into nu/nu recipients¹². The data summarized above suggest that at least a portion of these T cells recognize IgG2a^b-derived peptides as they can be specifically immunized or downregulated by IgG2a^b protein. To test this, we assessed the ability of BALB/c CD4⁺ Th₁ clones C1-6 and C1-15, which are specific for IgG2a^b (refs 13, 14), to transfer HSK into BALB/c nu/nu recipients (Fig. 2c). Recipients of C1-6 and C1-15 developed HSK, and this was associated with a substantial T-cell response to IgG2a^b in lymph nodes draining virally infected corneal beds at 10–14 days, consistent with homing of the clone to the site of infection (Fig. 3a). By contrast, recipients of BALB/c CD4⁺ Th₁ clones O3 (ref. 15), 8.7C11 and 8.7D8 (ref. 16), which are specific for peptides derived from ovalbumin and myelin basic protein, respectively, did not develop significant HSK (Fig. 2c).

We then investigated whether the pathogenic IgG2a^b-specific clones cross-recognized a self antigen expressed by corneal cells. The IgG2a^b-specific clones responded to murine corneal protein extracts but had no detectable response to identically prepared protein extracts from murine kidney and skin, suggesting that the antigen was tissue restricted (Fig. 3b). In contrast, the non-pathogenic CD4 clones O3, 8.7C11 and 8.7D8 had no detectable

reactivity to the same corneal extracts. Reactivity of the IgG2a^b-specific clones to corneal protein required CD28 ligation (which had no effect on the response of the control clones), suggesting that increased co-stimulatory activity associated with HSV-1 infection might enhance this response *in vivo*.

To define peptides that are derived from IgG2a^b and recognized by the two pathogenic clones, we compared the amino-acid sequence of IgG2a^b with IgG2a^a. Allotype-specific amino-acid differences are clustered in 5 distinct regions: 103–118; 215–230; 249–264; 270–285; and 292–308. IgG2a was also analysed using the TSites program¹⁷ to define residues that matched IA/I-E peptide motifs, in view of the finding that CD4 cells are necessary and sufficient for HSK transfer¹². Peptides 103–118, 249–264 and 292–308 contain potential IA motifs, and the first two sub-sequences also contain Claverie/Kourilsky motifs^{18,19}. Both IgG2a^b-specific clones were efficiently activated by peptide 292–308, but not by any of the other IgG2a^b peptides tested (Fig. 3b). Moreover, *in vivo* administration of soluble peptide 292–308, but not control peptides, two weeks before corneal inoculation of HSV-1 also resulted in resistance to HSK (Fig. 4a). In contrast, a 17-mer (301–315) that lacks the amino-terminal half of the putative IA motif (KLRVQKS) contained in the 292–308 sub-sequence had no stimulatory activity *in vitro* (Fig. 3b), and had no effect on HSK *in vivo* (Fig. 4b). This IgG2a^b-derived peptide is also recognized by T-cell clones that mediate IgG2a^b-specific suppression²⁰ and may represent a dominant self-peptide of IgG2a^b which determines the T-cell response to cells that display this isotype²¹. Finally, we investigated whether T cells from mice immunized to peptide 292–308 can transfer HSK susceptibility to syngenic nu/nu recipients, as they can for the intact IgG2a^b protein. Recipients of T cells immune to peptide 292–308, but not T cells from mice immunized with other peptides, developed HSK (Fig. 4b).

Sequestration of corneal tissue from the immune system and thymus is a consequence of the sparse supply of blood and lymphatics to this tissue²² and the limited transport of corneal antigens into regional lymphoid tissues^{23,24}. However, despite the poor access of corneal antigens to the developing thymus, clinical tolerance to corneal tissue is generally maintained throughout life. Inefficient presentation of corneal antigens to the small numbers of autoreactive T cells that enter this tissue may play a role in avoiding autoimmunity²². A second protective mechanism may depend on expression of shared peptides by corneal proteins and circulating immunoglobulin, resulting in T-cell tolerance to corneal self peptides. We suggest that peripheral tissues, such as cornea, may escape the penalties of immunological isolation by including peptide sub-sequences within tissue-specific proteins which crossreact with peptides that are derived from proteins with easy access to the thymus. We have been able to define an example of this mechanism because of a genetic polymorphism that affects one such shared T-cell epitope. Whether other genetic polymorphisms that affect resistance to autoimmune disease do so by this mechanism remains to be established. □

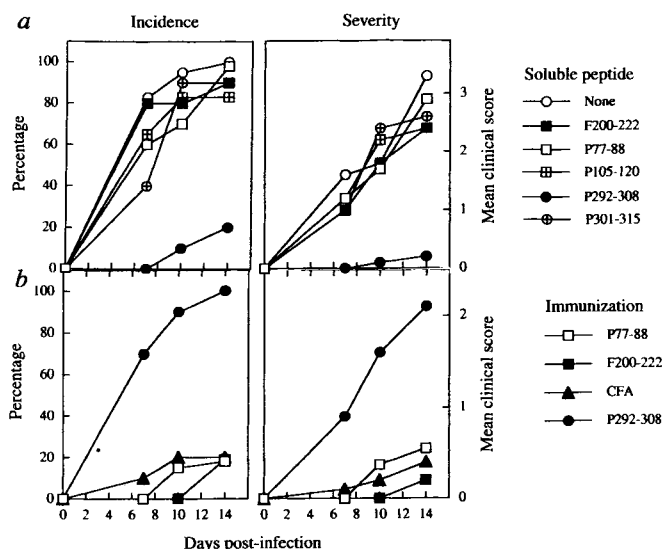


FIG. 4 Manipulation of HSK by IgG2a^b-derived peptidic fragments. a, Groups of 5–6 C.AL-20 mice were injected intravenously with 50 µg of the indicated synthetic peptides per mouse on days –14 and –7 before corneal HSV-1 infection. HSK was evaluated as described in Fig. 1a. b, C.AL-20 mice were immunized by tail-base injection with the indicated peptides (50 µg per mouse) emulsified in complete Freund's adjuvant. Two weeks later, purified T cells obtained from draining lymph nodes were used for adoptive transfer into BALB/c nu/nu recipients (5 × 10⁶ cells per host) immediately before HSV-1 corneal infection. HSK was measured as described in Fig. 1a.

Received 4 May; accepted 8 June 1995.

- Vidovic, D. & Matzinger, P. *Nature* **336**, 222–225 (1988).
- Matzinger, P. *Nature* **292**, 497–501 (1981).
- Metcalfe, J. F., Hamilton, D. S. & Reichert, R. W. *Infect. Immunity* **28**, 1164–1171 (1979).
- Russell, R. G., Nasaisse, M. P., Larson, H. S. & Rouse, B. T. *Invest. Ophthalmol. Vis. Sci.* **25**, 938–944 (1984).
- Mercadal, C. M., Bouley, D. M., DeStephano, D. & Rouse, B. T. *J. Virol.* **67**, 3404–3408 (1993).
- Dawson, C. & Togni, G. *Surv. Ophthalmol.* **21**, 11–29 (1976).
- Foster, C. S. *et al. Clin. Immun. Immunopathol.* **40**, 313–325 (1986).
- Garchon, H.-J. & Bach, J.-F. *Hum. Immun.* **32**, 1–30 (1991).
- Opremacak, E. M. *et al. Invest. Ophthalmol. Vis. Sci.* **29**, 749–754 (1988).
- DeWit, D. *et al. J. exp. Med.* **175**, 9–14 (1992).
- Metcalfe, J. F., Hamilton, D. S. & Reichert, R. W. *Infect. Immunity* **26**, 164–171 (1979).
- Akova, Y. A., Dutt, J., Rodriguez, A., Jabbur, N. & Foster, C. S. *Curr. Eye Res.* **12**, 1093–1101 (1993).
- Bikoff, E. & Eckhardt, L. A. *Eur. J. Immun.* **19**, 1903–1909 (1989).
- Bikoff, E. K. *Eur. J. Immun.* **21**, 1411–1417 (1991).
- Friedman, S., Silfcocks, D. & Cantor, H. *Immunogenetics* **26**, 193–203 (1987).
- Abromson-Leeman, S. *et al. J. Immun.* **154**, 388–398 (1995).

17. Feller, D. C. & de la Cruz, V. F. *Nature* **349**, 720–721 (1991).
18. Kourilsky, P. & Claverie, J.-M. *Adv. Immun.* **45**, 107–193 (1989).
19. Claverie, J.-M. *et al. Eur. J. Immun.* **18**, 1547–1553 (1988).
20. Bartnes, K. & Hannestad, K. *Eur. J. Immun.* **21**, 2365–2371 (1991).
21. Moudgil, K. D. & Sercarz, E. E. *J. exp. Med.* **178**, 2131–2138 (1993).
22. Niederkorn, J. Y. *Adv. Immun.* **48**, 191–226 (1990).
23. Ross, J., Callanan, D., Kunz, H. & Niederkorn, J. Y. *Invest. Ophthalmol. Vis. Sci.* **30**, 442 (1989).
24. Callanan, D., Peeler, J. & Niederkorn, J. Y. *Transplantation* **45**, 437–443 (1988).
25. Opremac, E. M., Rice, B. A., Wells, P. A. & Foster, C. S. *Invest. Ophthalmol. Vis. Sci.* **31**, 305–312 (1990).
26. Marshak-Rothstein, A., Siekevitz, M., Margolies, M. N., Mudgett-Hunter, M. & Geffer, M. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1120–1124 (1980).
27. Bikoff, E. & Birshtein, B. K. *J. Immun.* **137**, 28 (1986).
28. Alexander, R. J., Silverman, B. & Henley, W. L. *Expl Eye Res.* **32**, 205–216 (1981).

ACKNOWLEDGEMENTS. A.C.A. and Z.-S.Z. contributed equally to this study. This work was supported in part by the NIH and the Arthritis Foundation (to E.K.B., C.S.F. and H.C.). We thank A. Marshak-Rothstein for hybridoma proteins; S. Abromson-Leeman for T-cell clones; L. Gilmcher for a critical review; and A. Angel for assistance in the preparation of this manuscript.

First experimental transmission of fatal familial insomnia

Jun Tateishi*, Paul Brown†, Tetsuyuki Kitamoto*, Zahirul M. Hoque*, Raymond Roos§, Robert Wollman||, Larisa Cervenáková‡ & D. Carleton Gajdusek‡

* Department of Neuropathology, Neurological Institute, Kyushu University, Fukuoka 812, Japan

† Laboratory of CNS Studies, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892, USA

§ Departments of Neurology§ and Pathology||, University of Chicago Medical Center, Chicago, Illinois 60637, USA

ORIGINALLY described by Lugaresi *et al.* in 1986 (ref. 1), fatal familial insomnia (FFI) is a rare inherited neurological disease characterized by the subacute progression of intractable insomnia and other autonomic abnormalities, cerebellar and pyramidal signs, myoclonus and dementia; neuropathologically, the major feature is severe neuronal loss with associated gliosis in the ventral and mediodorsal thalamic nuclei. The disease has been related to the group of spongiform encephalopathies by virtue of the presence of low levels of proteinase-resistant amyloid protein (PrP^{res}) in the brain^{2–4}, and of a pathogenic single-allele mutation at codon 178 of the *PRNP* gene that encodes PrP^{res} (refs 2, 5). Here we report the successful transmission of the disease to experimental animals, placing FFI within the group of infectious cerebral amyloidoses.

We established the diagnosis of fatal familial insomnia in a patient whose immediate relatives had no knowledge of any familial occurrence of the disease, but who was later discovered to have ancestral ties to a previously reported American FFI family⁶. The illness began with episodic sensory, motor and visual complaints and thereafter followed a fairly typical clinical course that included intractable insomnia, with characteristic thalamic pathology, low levels of PrP^{res} in the brain, and the FFI genetic 'signature' *PRNP* genotype (codon 178^{Asn}/codon 129^{Met}); like other affected members of the distantly related branch of his family, he also had a 24-base-pair (bp) deletion between codons 51 and 91 (ref. 7).

A piece of the patient's thalamus was homogenized as a tissue suspension (200 mg ml⁻¹) in PBS (pH 7.4), from which 0.02-ml aliquots were inoculated intracerebrally into each of 19 NZW strain mice. As a control, 24 mice were given parallel inoculations of a tissue homogenate of frontal cortex from a demented patient with an identical 24-bp deletion but no PrP^{res} in the brain. All mice were observed for two years, with symptomatic illnesses noted, and all dead mice (except one that died from inoculation trauma) were examined neuropathologically and histochemically. Sections used for histochemical examination were pretreated with a dilute solution of HCl followed by autoclaving

TABLE 1 Experimental mortality in mice inoculated with brain homogenates from a patient with FFI and a control patient with a non-spongiform dementia

FFI			Control		
Days after inoculation	Neuro-pathology	PrP ^{res}	Days after inoculation	Neuro-pathology	PrP ^{res}
218	–	+	106	–	–
397	+	+	203	–	–
421	+	+	403	–	–
425	+	+	417	–	–
425	+	+	618	–†	–
428	+	+			
429	+	+			
432	+	+			
432	+	+			
432	+	+			
453	+	+			
481	+	+			
491	+	+			
492	+	+			
502	+	+			
506	+	+			
>620	3 mice still alive and well		19 mice still alive and well		

PrP^{res} denotes 'proteinase-resistant protein'. Mean incubation period $\pm$ s.d. until death of mice inoculated with FFI was 455 ± 35 days (not including the mouse that died on day 218).

* Figure 1a.

† Figure 1b.

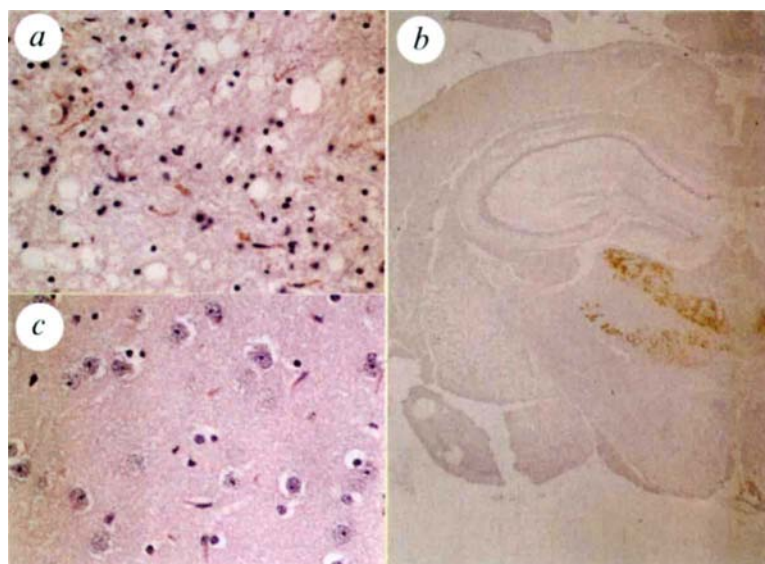
‡ Figure 1c.

and then successively incubated with rabbit antibodies to PrP^{res}, swine anti-rabbit IgG, and a peroxidase–rabbit anti-peroxidase complex; colour was developed with diaminobenzidine tetrahydrochloride⁸.

In the group of 19 mice inoculated with the FFI brain homogenate, one mouse died shortly after inoculation and was discarded. Of the remaining 18 mice, one was found dead on day 218, 14 mice developed typical signs of spongiform encephalopathy and died between days 397 and 506, and 3 mice are still alive without signs of disease 620 days after inoculation (Table 1). The mouse dying on day 218 had no neuropathological changes but did show a minimal deposition of PrP^{res} in the thalamus. The 14 symptomatic mice dying between days 397 and 506 all had characteristic spongiform change, astrocytic proliferation, and neuronal loss (Fig. 1a) as well as PrP^{res} deposition (Fig. 1b). These deposits were more pronounced in the thalamus than is usual in primary isolations from other types of human spongiform encephalopathy⁹; however, in all other respects, FFI produced clinical and neuropathological features that were indistinguishable from those produced by brain homogenates from patients with Creutzfeldt–Jakob disease (CJD). In the group of mice inoculated with the control homogenate, 5 died or were killed for study at various times during the 620-day observation period, showing neither spongiform change (Fig. 1c) nor PrP^{res} deposition; 19 mice remain alive and healthy.

In laboratories working with strains of transmissible agents that have been adapted to animals, a question always arises as to whether primary transmissions might be an artefact of laboratory cross-contamination. In the present case, cross-contaminating infection was highly unlikely because: (1) none of the control mice became ill; (2) incubation periods in the mice inoculated with the patient's brain homogenate were in line with the usual 400–600-day incubation periods of primary human isolations rather than with the shorter 120–150-day incubation periods that typify infection with mouse-adapted strains¹⁰; and (3) none of the affected mice had PrP^{res} deposits in the follicular dendritic cells of the spleen, which is a characteristic finding in mice infected intracerebrally with mouse-adapted strains after short incubation periods¹¹, and which has also been seen after incubation periods of up to 625 days in mice inoculated with low-

FIG. 1 a, Photomicrograph showing severe gliosis, loss of nerve cells and mild vacuolation in the thalamus of a symptomatic mouse that 432 days earlier had been inoculated with a suspension (200 mg ml^{-1}) of thalamus from a patient with fatal familial insomnia (haematoxylin and eosin; original magnification, $\times 390$). b, Immunostaining of proteinase-resistant protein (PrP^{res}) in the thalamus of the same mouse. The hydrolytically autoclaved formalin-fixed section was incubated successively with a rabbit anti- PrP^{res} antibody, swine anti-rabbit IgG, and a peroxidase–rabbit anti-peroxidase complex; colour was developed with diaminobenzidine tetrahydrochloride (original magnification, $\times 22$). c, No histological abnormality in the thalamus of an asymptomatic mouse that died 618 days after inoculation of control brain suspension (haematoxylin and eosin; original magnification, $\times 390$).



titre intestinal tissue from animals infected with mouse-adapted strains (T.K. and J.T., unpublished results).

Previous failures to transmit FFI to experimental animals may have resulted from the use of inocula prepared from pathologically uninvolved neocortex containing little or no infection-specific PrP^{res} . In a study of 37 patients with various forms of spongiform encephalopathy (sporadic, iatrogenic and familial), we found an overall positive correlation between levels of PrP^{res} (as demonstrated in western blots of brain tissue extracts) and transmission rates to primates: PrP^{res} levels in FFI brains were lower than in any other form of disease, and also the most focally restricted⁴. Our patient had more PrP^{res} in the temporal lobe than in the thalamus, but amounts were minimal in all areas of the brain⁷. We have inoculated separate specimens prepared from neocortex, basal ganglia and thalamus from this patient, as well as from patients in two other (unrelated) FFI families in order to clarify the issue, but it may be some years before the results are known (so far, 1–2 years after inoculation, all animals remain well).

The relationship of FFI to other human spongiform encephalopathies has been controversial because of its distinctive clinicopathological features, and its failure until now to transmit disease when inoculated into experimental animals (from three different members of a single family)^{1,2}. However, at least one affected individual in all reported FFI families has suffered an illness resembling CJD with clear-cut spongiform encephalopathy, and PrP^{res} is consistently detectable in detergent extracts of FFI brains, albeit in comparatively low concentration^{2–4}. Our demonstration of its experimental transmissibility closes discussion of its proper classification: it is a genotypically and phenotypically distinctive subtype of the transmissible spongiform encephalopathy disease group. □

Essential role for ZAP-70 in both positive and negative selection of thymocytes

Izumi Negishi*, Noboru Motoyama, Kei-ichi Nakayama*, Kelko Nakayama*, Satoru Senju, Shigetsugu Hatakeyama, Qing Zhang, Andrew C. Chan† & Dennis Y. Loh*‡

Howard Hughes Medical Institute, Departments of Medicine, Genetics and Molecular Microbiology, * Division of Rheumatology, Departments of Medicine and Pathology, Washington University School of Medicine, St Louis, Missouri 63110, USA

DURING thymic development, T cells that can recognize foreign antigen in association with self major histocompatibility complex (MHC) are selected for survival (positive selection) and autoreactive T cells are eliminated (negative selection). Both of these selective events are mediated by interaction between the T-cell receptor (TCR) and the peptide–MHC complex¹. But the signalling pathways that lead to cell survival or to cell death are still unclear. ZAP-70 is a protein tyrosine kinase (PTK) that is associated with the TCR signalling subunits (CD3 and ζ) and is expressed in T cells and natural killer cells². It has been shown that ZAP-70 plays a crucial role in T-cell activation^{2–5} and development^{6–8}. Here we show that mice lacking ZAP-70 had neither CD4 nor CD8 single-positive T cells, but human ZAP-70 reconstituted both CD4 and CD8 single-positive populations. Moreover, ZAP-70^{−/−} thymocytes were not deleted by peptide antigens. Natural killer cell function was intact in the absence of ZAP-70. These data suggest that ZAP-70 is a central signalling molecule during thymic selection for CD4 and CD8 lineage.

ZAP-70-deficient mice were generated by deleting the entire gene segment encoding ZAP-70 protein (Fig. 1a). ZAP-70 protein could not be detected in thymocytes from ZAP-70^{−/−} mice by western blot analysis² (data not shown). ZAP-70^{−/−} mice appeared healthy and fertile under specific pathogen free (SPF) conditions until at least 10 months of age.

Received 26 April; accepted 12 June 1995.

1. Lugaresi, E. et al. *New Engl. J. Med.* **315**, 997–1003 (1986).
2. Medori, R. et al. *New Engl. J. Med.* **326**, 444–449 (1992).
3. Monari, L. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 2839–2842 (1994).
4. Brown, P. et al. *Ann. Neurol.* **38**, 245–253 (1995).
5. Goldfarb, L. G. et al. *Science* **258**, 806–808 (1992).
6. Bosque, P. J., Vnencak-Jones, C. L., Johnson, M. D., Whitlock, J. A. & McLean, M. J. *Neurology* **42**, 1864–1870 (1992).
7. Reder, A. T. et al. *Neurology* **45**, 1068–1075 (1995).
8. Kitamoto, T. et al. *Am. J. Path.* **140**, 1285–1294 (1992).
9. Muramoto, T., Kitamoto, T., Tateishi, J. & Goto, I. *Brain Res.* **599**, 309–316 (1992).
10. Tateishi, J. et al. *Acta neuropath.* **53**, 161–163 (1981).
11. Muramoto, T., Kitamoto, T., Tateishi, J. & Goto, I. *Am. J. Path.* **140**, 1411–1420 (1992).
12. Little, B., Brown, P. W., Rodgers-Johnson, P., Perl, D. P. & Gajdusek, D. C. *Ann. Neurol.* **20**, 231–239 (1986).

* Present address: Department of Biology, Nippon Roche Research Center, 200 Kajiwa, Kamakura Kanagawa 247, Japan.

‡ To whom correspondence should be addressed.

A rare human severe combined immunodeficiency (SCID) is associated with genetic mutations in the ZAP-70 gene⁶⁻⁸. These patients have no CD8 single-positive (SP) T cells but have normal or raised numbers of non-functional CD4 SP T cells. In ZAP-70^{-/-} mice, thymocyte number ($1.7 \times 10^8 \pm 7.1 \times 10^7$, $n = 7$) was comparable to those from wild-type littermates ($1.9 \times 10^8 \pm 3.5 \times 10^7$, $n = 5$). However, unlike the human SCID cases, neither CD4 nor CD8 SP thymocytes, and very few $\alpha\beta$ TCR^{high} mature thymocytes, were detected in ZAP-70^{-/-} mice by flow cytometric analysis (Fig. 1b). Cell size and the expression levels of CD3, $\alpha\beta$ -TCR, CD24, CD25 and CD44 were indistinguishable between wild-type and ZAP-70^{-/-} double-positive (DP) thymocytes (data not shown). Splenocyte

number ($8.4 \times 10^7 \pm 2.0 \times 10^7$, $n = 7$) was reduced by about 30% in ZAP-70^{-/-} mice compared to wild-type mice ($1.3 \times 10^8 \pm 3.7 \times 10^7$, $n = 5$). No $\alpha\beta$ -TCR⁺ T cells were present and the proportion of B220⁺ cells was relatively increased in the ZAP-70^{-/-} spleen (Fig. 1b). Histological study of ZAP-70^{-/-} thymus revealed a gross expansion of the cortex and a rudimentary medulla (Fig. 1c), consistent with previous reports suggesting that mature T cells are required for the formation and maintenance of thymic medulla⁹⁻¹¹. No increase of intracellular free calcium ($[Ca^{2+}]_i$) was detected in ZAP-70^{-/-} thymocytes after TCR crosslinking with an anti-CD3 monoclonal antibody (Fig. 1d).

These phenotypic and functional defects of ZAP-70^{-/-} T cells were compensated with human wild-type ZAP-70 (Fig. 2).

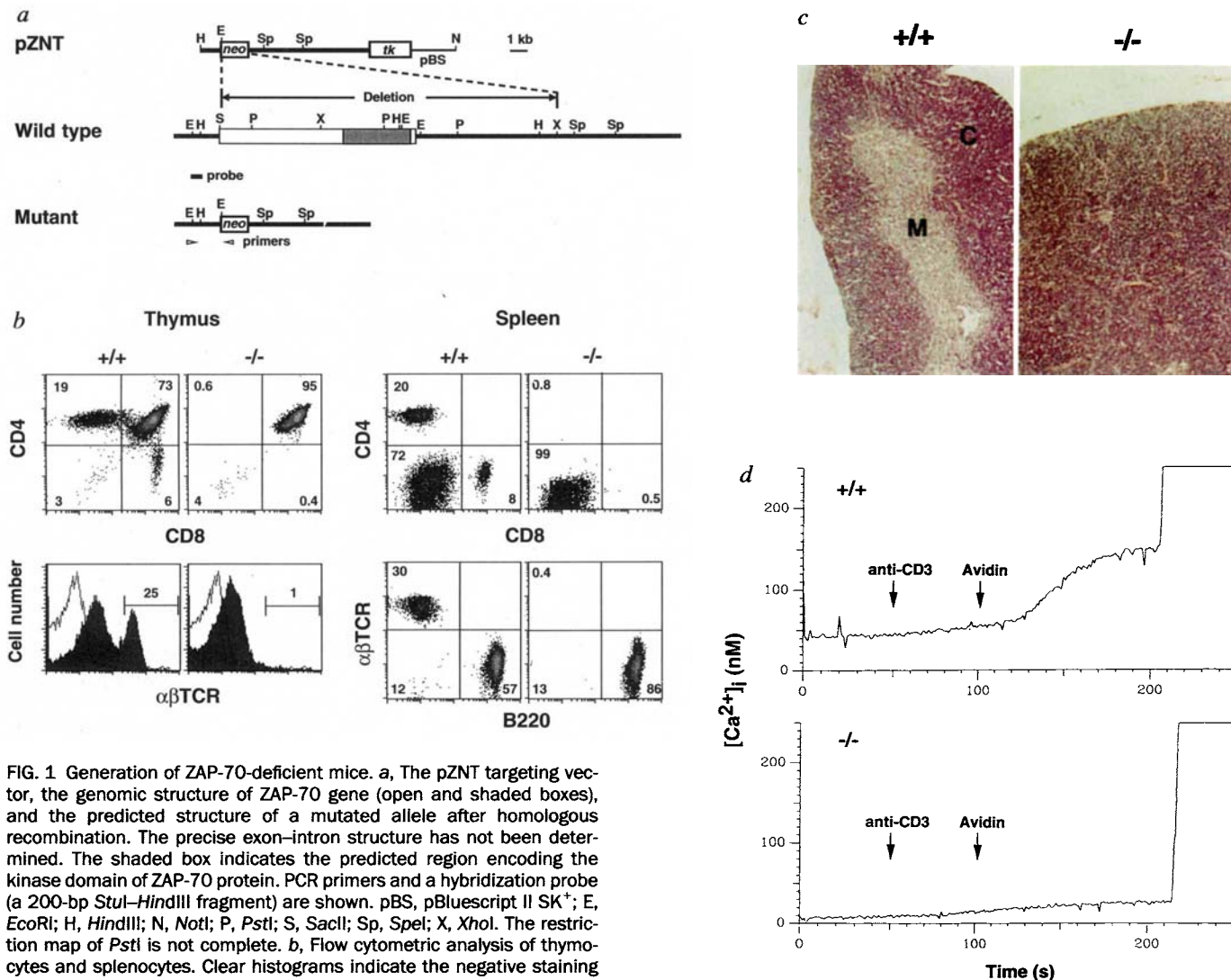


FIG. 1. Generation of ZAP-70-deficient mice. **a**, The pZNT targeting vector, the genomic structure of ZAP-70 gene (open and shaded boxes), and the predicted structure of a mutated allele after homologous recombination. The precise exon-intron structure has not been determined. The shaded box indicates the predicted region encoding the kinase domain of ZAP-70 protein. PCR primers and a hybridization probe (a 200-bp *StuI*-*HindIII* fragment) are shown. pBS, pBluescript II SK⁺; E, *EcoRI*; H, *HindIII*; N, *NotI*; P, *PstI*; S, *SacII*; Sp, *SpeI*; X, *XhoI*. The restriction map of *PstI* is not complete. **b**, Flow cytometric analysis of thymocytes and splenocytes. Clear histograms indicate the negative staining control. Percentages of each population are indicated. **c**, Histological examination of the thymus ($\times 100$). The cortex (C) and medulla (M) are indicated in the ZAP-70^{+/+} thymus. The thymuses were fixed in 10% buffered formalin and embedded in paraffin. Sections were stained with a haematoxylin and eosin. **d**, Lack of mobilization of $[Ca^{2+}]_i$ in the ZAP-70^{-/-} thymocytes.

METHODS. **a**, Genomic DNA corresponding to the ZAP-70 locus was isolated from a library of 129/Sv mouse DNA (Stratagene). The *SacII*-*XhoI* fragment (~20 kb), containing most of the ZAP-70 coding region except for 14 bp from the initiation codon, was deleted and replaced with a PGK-neo-polyadenylate [poly(A)] cassette. The pZNT targeting vector contains 1.4 kb of homology 5', and 8.0 kb 3' of the drug-resistant gene and a PGK-tk-poly(A) cassette. The linearized pZNT was transfected into E14 embryonic stem (ES) cells. Mutant ES cells were injected into C57BL/6 (B6) blastocysts. Male chimaeric mice were mated with B6 females and heterozygous mutant (ZAP-70^{+/+}) mice were

intercrossed to obtain homozygous mutant (ZAP-70^{-/-}) mice. Maintenance, transfection and selection were carried out as previously described²¹. For flow cytometric analysis, thymocytes (1×10^6) were stained with anti-CD4-phycoerythrin (PE), anti-CD8-fluorescein isothiocyanate (FITC) and anti- $\alpha\beta$ TCR-biotin, followed by Streptavidin-Cy-Chrome (PharMingen). Spleen cells (1×10^6) were stained with anti-CD4-PE and anti-CD8-FITC, or anti- $\alpha\beta$ TCR-PE and anti-B220-FITC; 2×10^4 viable cells were analysed with the CELLQuest program (Becton-Dickinson). All antibodies were purchased from PharMingen. Five ZAP-70^{+/+} or ZAP-70^{-/-} mice at the age of 4–8 weeks were analysed. To measure $[Ca^{2+}]_i$, thymocytes were loaded with Fura-2 and analysed by spectrofluorimetry as previously described⁸. Cells were stimulated with a biotinylated anti-CD3 ϵ antibody (145-2C11, $5 \mu\text{g ml}^{-1}$) and crosslinked with avidin ($5 \mu\text{g ml}^{-1}$). Cells were also stimulated with ionomycin to ensure that cells were equivalently loaded with Fura-2.

Expression of human wild-type ZAP-70 in ZAP-70^{-/-} thymocytes reconstituted the development of CD4 and CD8 SP T cells (Fig. 2a), as well as TCR-mediated induction of the early activation marker, CD69 (ref. 12) (Fig. 2b). Taken together, these data indicate that ZAP-70 is indispensable for T-cell development and activation, but not for B-cell development.

There are a few possibilities to explain the presence of CD4 SP T cells in ZAP-70^{-/-} patients. Syk, which is structurally homologous to ZAP-70 and is also expressed in thymocytes^{3,13}, may compensate for the absence of ZAP-70 function and mediate the positive selection of only CD4 SP thymocytes⁶⁻⁸. Alternatively, the TCR signalling pathways in humans and mice may differ. But this appears unlikely, because both CD4 and CD8 SP populations were reconstituted with the human wild-type ZAP-70 (Fig. 2a). Three different mutations have been identified in the ZAP-70 gene of the SCID patients: a 9-base-pair (bp) insertion, a 13-bp deletion, or a point mutation, all within the kinase domain of ZAP-70 (refs 6-8). Our observations in mice are from a complete absence of ZAP-70 protein (Fig. 1a). Thus, it is also possible that the mutant ZAP-70 proteins, which are not detected by immunoblotting or an *in vitro* kinase assay, may be expressed at a low level and mediate the positive selection of only CD4 SP cells in the SCID patients. Alternatively, these

patients may have developed an increase in the level of other protein tyrosine kinases, such as Syk, to compensate partially for the absence of ZAP-70.

Thymocytes bearing an autoreactive TCR are deleted by apoptosis within the thymus during development¹. Anti-CD3 treatment can induce apoptosis of thymocytes and mimic negative selection *in vitro*¹⁴. The ability of ZAP-70^{+/+} and ZAP-70^{-/-} thymocytes to undergo apoptosis was evaluated by *in vitro* treatment of cells with dexamethasone (Dex) or an anti-CD3 antibody (Fig. 3a). Both ZAP-70^{+/+} and ZAP-70^{-/-} thymocytes were deleted by Dex. In contrast, although anti-CD3 antibody treatment resulted in death of >50% of the ZAP-70^{+/+} DP thymocytes, no apoptosis of ZAP-70^{-/-} thymocytes was detected. Thymocytes from DO10 TCR transgenic mice are deleted by chicken ovalbumin (cOVA) peptide (323-339) at the DP stage *in vivo*¹⁵ and *in vitro*¹⁶. To evaluate further the role of ZAP-70 in negative selection, ZAP-70^{-/-} mice were crossed with DO10 mice to obtain DO10 ZAP-70^{-/-} mice. Although most of the DP thymocytes from DO10 ZAP-70^{+/+} mice were deleted by cOVA-(323-339), DO10 ZAP-70^{-/-} thymocytes were resistant to this deletion (Fig. 3b). The thymocytes of both genotypes were not deleted by medium alone or the cOVA-(324-334) control peptide. These results suggest that ZAP-70 is also indispensable for peptide antigen-driven negative selection.

ZAP-70 is expressed not only in T cells but also in natural killer (NK) cells². NK cells exert natural cytotoxicity against some tumour cells and antibody-dependent cellular cytotoxicity (ADCC) against antibody-coated target cells. ADCC is mediated by the low-affinity IgG receptor, CD16 (FcγRIIIA), which is associated with the TCR-ζ and FcεRIγ chains¹⁷. ZAP-70 has been proposed to be involved in the activation pathway of ADCC^{18,19}. NK cell activity of splenocytes was assessed using

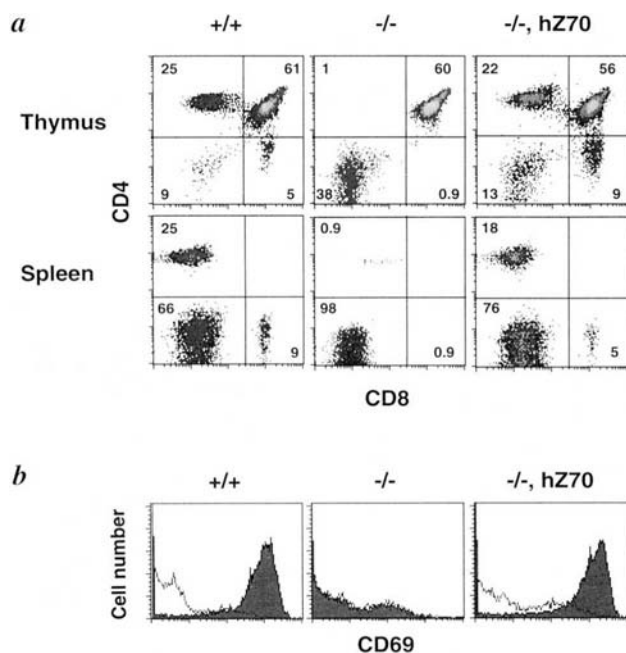


FIG. 2 Reconstitution of CD4 and CD8 SP T cells. a, ZAP-70^{+/+} (+/+), ZAP-70^{-/-} (-/-), or human wild-type ZAP-70-expressing ZAP-70^{-/-} (-/-, hZ70) thymocytes and splenocytes were analysed for the expression of CD4 and CD8. The increase of the double-negative (DN) thymocytes is due to the mixture of the cells from RAG-2^{-/-} blastocysts and ES cells. b, Induction of CD69 expression on peripheral T cells. Lymph node cells (5×10^5) were cultured in 96-well plates with (shaded histograms) or without (clear histograms) an anti-CD3ε antibody (145-2C11, $1 \mu\text{g ml}^{-1}$) for 16 h. Cells were stained with anti-CD69-biotin, followed by streptavidin-PE.

METHODS. The pZHT targeting vector containing a PGK-hyg-poly(A) cassette instead of a PGK-neo-poly(A) cassette was constructed as described in Fig. 1. To obtain ZAP-70^{-/-} ES cells, the pZHT was transfected into ZAP-70^{+/+} ES cells²¹. Then human wild-type ZAP-70 cDNA, which was inserted into the *lck-hGH* vector²², was cotransfected with a puromycin-resistant gene into ZAP-70^{-/-} ES cells²¹. ZAP-70^{+/+}, ZAP-70^{-/-} ES cells, and ZAP-70^{-/-} ES cells carrying human wild-type ZAP-70 gene were injected into RAG-2^{-/-} blastocysts²³. Thymocytes and splenocytes from RAG-2 chimaeric mice were analysed by a flow cytometer as described in Fig. 1.

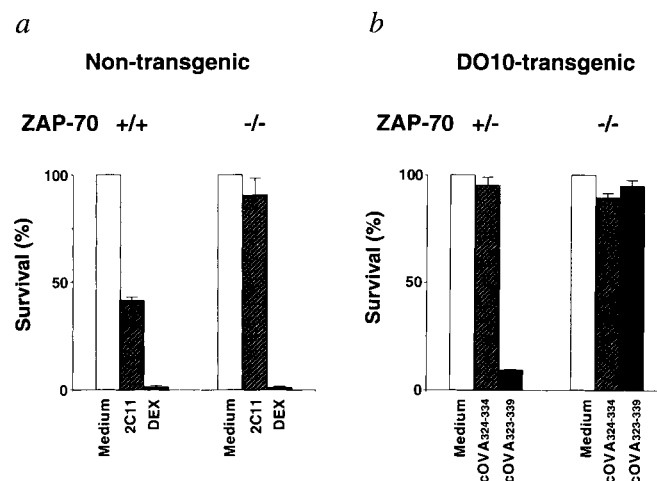
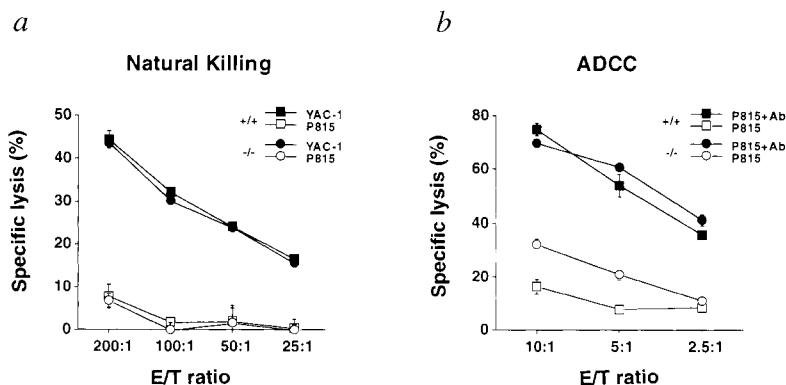


FIG. 3 *In vitro* negative selection of thymocytes. a, Anti-CD3-induced apoptosis of thymocytes. Thymocytes from non-transgenic ZAP-70^{+/+} or ZAP-70^{-/-} mice were cultured with medium alone, an anti-CD3ε antibody (145-2C11, $1 \mu\text{g ml}^{-1}$) or Dex (100 nM). b, Peptide-induced apoptosis of thymocytes. Thymocytes from DO10 ZAP-70^{+/+} or DO10 ZAP-70^{-/-} mice were cultured with medium alone, cOVA-(324-334) ($1 \mu\text{M}$) or cOVA-(323-339) ($1 \mu\text{M}$). These figures represent four and two independent experiments in a and b, respectively.

METHODS. Thymocytes (1×10^5) were cultured with or without each additive in the presence of 1×10^4 A20 cells (B-cell lymphoma) for 20 h in a V-bottomed 96-well plate (Costar)¹⁶. Cells were stained with anti-CD4-PE and anti-CD8-FITC antibodies for a flow cytometric analysis. Cell viability was determined by propidium iodide staining and by forward scatter (FSC)/side scatter (SSC) parameters. Per cent cell survival was calculated as (number of viable DP cells cultured with each additive/number of viable DP cells cultured in medium alone) $\times 100$. A20 cells were gated out by FSC/SSC parameters.

FIG. 4 NK cell function of splenocytes from ZAP-70^{+/+} (+/+) or ZAP-70^{-/-} (-/-) mice. **a**, Natural killing activity against YAC-1 or P815 target cells. **b**, ADCC activity against P815 cells in the presence or absence of anti-P815 antiserum.

METHODS. Three mice of each genotype were analysed in natural killing and ADCC assays using standard ⁵¹Cr-release assay. For the natural killing assay, spleen cells depleted of red blood cells were cultured at the indicated effector/target (E/T) ratio with 5,000 ⁵¹Cr-labelled YAC-1 or P815 cells for 4 h²⁴. For the ADCC assay, nylon wool-passed spleen cells were stimulated with 1,000 U ml⁻¹ mouse rIL-2 (PharMingen) for 7 days and adherent cells were used as effector cells²⁵. Effector cells were cultured at indicated E/T ratio with 5,000 ⁵¹Cr-labelled P815 cells in the presence or absence of anti-P815 antiserum. Mice were injected intraperitoneally with 200 µg per mouse of INF inducer tilorone (T-8014, Sigma) 24 h before collecting spleen cells²⁴.



both a natural killing and an ADCC assay. The splenocytes from both ZAP-70^{+/+} and ZAP-70^{-/-} mice caused cytolysis of NK-sensitive YAC-1 cells but not of NK-insensitive P815 tumour cells (Fig. 4a). Unexpectedly, the ZAP-70^{-/-} splenocytes could lyse P815 cells as well as wild-type cells in the presence of anti-P815 antiserum (Fig. 4b). These findings demonstrate that ZAP-70 is dispensable for NK cell development and function.

ZAP-70 is expressed in all thymocyte subsets¹³. In addition, ZAP-70 is constitutively associated with a tyrosine-phosphorylated ζ-chain in DP thymocytes and TCR ligation induces tyrosine-phosphorylation of ZAP-70 (ref. 20). We have shown that ZAP-70^{-/-} thymocytes expressing αβ-TCR are confined to the DP state just before thymic selection (Fig. 1b). It is likely that TCR engagement of DP thymocytes activates ZAP-70 for final maturation from DP to SP cells. Our data strongly suggest that ZAP-70 stands at the crossroads of the signalling pathways to determine the fate of thymocytes for both CD4 and CD8 lineage. It is of great importance to explore putative signalling molecules that associate with ZAP-70 and to understand the signals mediated by ZAP-70 that lead to either positive or negative selection. □

CO₂ fixation and photoevolution of H₂ and O₂ in a mutant of *Chlamydomonas* lacking photosystem I

E. Greenbaum*, J. W. Lee*, C. V. Tevault*, S. L. Blankinship* & L. J. Mets†

* Chemical Technology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6194, USA

† Department of Molecular Genetics and Cell Biology, University of Chicago, 1103 E. 57th Street, Chicago, Illinois 60637, USA

ALTHOUGH mutant B4 of *Chlamydomonas reinhardtii* lacks photosystem I (PS I), it is capable of photoautotrophic assimilation of atmospheric carbon dioxide and sustained simultaneous photoevolution of molecular oxygen and hydrogen. Here we report that at saturating light intensities, carbon dioxide reduction is stable under anaerobiosis but unstable in air. At lower light intensities, carbon dioxide reduction is stable in both atmospheres. The data indicate that PS I is not necessary for autotrophic photosynthesis. One interpretation of these results is that oxygenic photosynthesis developed as a single-light-reaction process, presumably from a bacterium with a phaeophytin-quinone reaction centre, but became unstable as oxygen in the Earth's atmosphere accumulated. PS I was the second light reaction, added to confer stability in oxygen-containing atmospheres. Viewed from this perspective, the well-known Z scheme of modern photosynthesis is seen as a specialized adaptation for performing low-potential reductive photochemistry in oxygen-containing atmospheres, but is not an irreducible necessity for satisfying the thermodynamic and mechanistic requirements of carbon dioxide photoreduction using water as the source of reductant.

Mutant B4 of *Chlamydomonas reinhardtii* was obtained after metronidazole enrichment described in detail elsewhere¹. Analysis of the algal cultures before and after the experiments indicated that this mutant had a chlorophyll-fluorescence induction transient characteristic of PS-I-deficient strains² and lacked the reaction centre core (CPI) complex in lithium dodecyl sulphate (LDS) electrophoresis at 4 °C. Complete absence of P700 in B4 samples was also confirmed by electron paramagnetic resonance and by absorbance spectroscopic measurements. Genetic analysis indicated that the B4 mutation is carried in a nuclear gene and is unable to assemble the functional message for the *psaA* gene encoding one of the two major apoproteins of the PS I

Received 17 March; accepted 16 June 1995.

- Robey, E. & Fowlkes, B. J. *A. Rev. Immun.* **12**, 675–705 (1994).
- Chan, A. C., Iwashima, M., Turck, C. W. & Weiss, A. *Cell* **71**, 649–662 (1992).
- Chan, A. C., Desai, D. M. & Weiss, A. *A. Rev. Immun.* **12**, 555–592 (1994).
- Sloan-Lancaster, J., Shaw, A. S., Rothbard, J. B. & Allen, P. M. *Cell* **79**, 913–922 (1994).
- Madrenas, J. et al. *Science* **267**, 515–518 (1994).
- Arpaia, E., Shahar, M., Dadi, H., Cohen, A. & Roifman, C. M. *Cell* **76**, 947–958 (1994).
- Elder, M. E. et al. *Science* **264**, 1596–1599 (1994).
- Chan, A. C. et al. *Science* **264**, 1599–1601 (1994).
- Philpott, K. L. et al. *Science* **256**, 1448–1452 (1992).
- Surh, C. D., Ernst, B. & Sprent, J. *J. exp. Med.* **176**, 611–616 (1992).
- van Ewijk, W., Shores, E. W. & Singer, A. *Immun. Today* **15**, 214–217 (1994).
- Testi, R., D'Ambrosio, D., De Maria, R. & Santoni, A. *Immun. Today* **15**, 479–483 (1994).
- Chan, A. C. et al. *J. Immun.* **152**, 4758–4766 (1994).
- Smith, C. A., Williams, G. T., Kingston, R., Jenkinson, E. J. & Owen, J. J. T. *Nature* **337**, 181–184 (1989).
- Murphy, K. M., Heimberger, A. B. & Loh, D. Y. *Science* **250**, 1720–1723 (1990).
- Iwabuchi, K. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 9000–9004 (1992).
- Ravetch, J. V. *Cell* **78**, 553–560 (1994).
- Vivier, E. et al. *Eur. J. Immun.* **23**, 1872–1876 (1993).
- Cone, J. C., Lu, Y., Trevillyan, J. M., Bjornndahl, J. M. & Phillips, C. A. *Eur. J. Immun.* **23**, 2488–2497 (1993).
- van Oers, N. S. C., Killeen, N. & Weiss, A. *Immunity* **1**, 675–685 (1994).
- Nakayama, K.-i. et al. *Science* **263**, 1131–1133 (1994).
- Chaffin, K. E. et al. *EMBO J.* **9**, 3821–3829 (1990).
- Chen, J., Lansford, R., Stewart, V., Young, F. & Alt, F. W. *Proc. natn. Acad. Sci. U.S.A.* **90**, 4528–4532 (1993).
- Mombaerts, P. et al. *Int. Immun.* **6**, 1061–1070 (1994).
- Gurji, Y., Vujanovic, N. L., Hiserodt, J. C., Herberman, R. B. & Gorelik, E. *J. Immun.* **142**, 1748–1754, (1989).

ACKNOWLEDGEMENTS. We thank F. W. Alt for RAG-2^{-/-} mice; K. Iwabuchi for an anti-CD3ε (145-2C11) mAb; M. W. Olszowy for YAC-1, P815 cells, an anti-P815 antiserum and technical advice; H. Rayburn and G.-H. Kong for technical advice; J. D. Alvarez and I. T. Chan for critical reading of the manuscript; P. J. Lucas and F. Wang for discussion. Supported by the Howard Hughes Medical Institute (D.Y.L. and A.C.C.) and by the Japan Society for the Promotion of Science (N.M.).

complex. This property has been described previously³ for several other nuclear-gene mutants of *C. reinhardtii* that evidently encode components essential for the *trans*-splicing of this tripartite message. In the absence of the *psaA* message, the cells were unable to synthesize any PS I complexes.

Figure 1a demonstrates simultaneous non-saturated photoevolution of molecular hydrogen and oxygen from mutant B4 in an atmosphere of 363 p.p.m. CO₂ in helium; Fig. 1b contains data for the simultaneous photoevolution of hydrogen and oxygen in pure helium. Fig. 1c, d contain corresponding data for wild-type *Chlamydomonas* 137c. Mutant B4 was grown photoheterotrophically on minimal growth medium with acetate. The reaction itself was performed in minimal growth medium by successive washings to eliminate all soluble acetate. (Preliminary experiments indicate that B4 is capable of growing photoautotrophically at a light intensity of 50 W m⁻².) Data for mutant B4 and the control wild-type 137c were obtained simultaneously in a dual-photoreaction chamber flow system illustrated in Fig. 2. To maintain anaerobiosis, even during photosynthetic oxygen production, the algal suspension was continuously purged with O₂-free carrier gas of ultra-high purity at 50 ml min⁻¹. Therefore regular respiration or chloroplast respiration do not occur in the experiments performed under anaerobiosis, either in helium or in CO₂ in helium. The experimental technique for measuring the

simultaneous photoevolution of oxygen and hydrogen has been previously described⁴. Even though B4 totally lacks PS I, sustained simultaneous photoevolution of hydrogen and oxygen was observed, demonstrating that a PS-II-driven reaction can span the potential difference between water oxidation/oxygen evolution and proton reduction/hydrogen evolution. Because the steady-state flow-stream concentrations (<10 p.p.m.) of hydrogen and oxygen are far from the standard state, these data place an upper limit on the potential spanned by the PS II light reaction that is less than the standard value. The reversible equilibrium thermodynamic potential for these two reactions is 1.23 V (ref. 5). Of course, because the reactions do not occur reversibly at thermodynamic equilibrium, the actual potential must be higher, by about 0.2–0.4 V (ref. 6).

In Fig. 1a (mutant) and c (wild type), for which CO₂ was present in the carrier gas, it can be seen that there is an initial surge of hydrogen that peaks and then declines to zero while the light is still on. This hydrogen photoevolution pattern for wild-type *Chlamydomonas* has been previously analysed⁷. Briefly, when the algae have been dark-adapted, the Calvin–Benson cycle is shut off. Consequently, upon illumination all reducing equivalents are shunted through the ferredoxin/hydrogenase pathway. As the Calvin–Benson cycle is activated by light, it becomes the exclusive sink for reducing equivalents. The

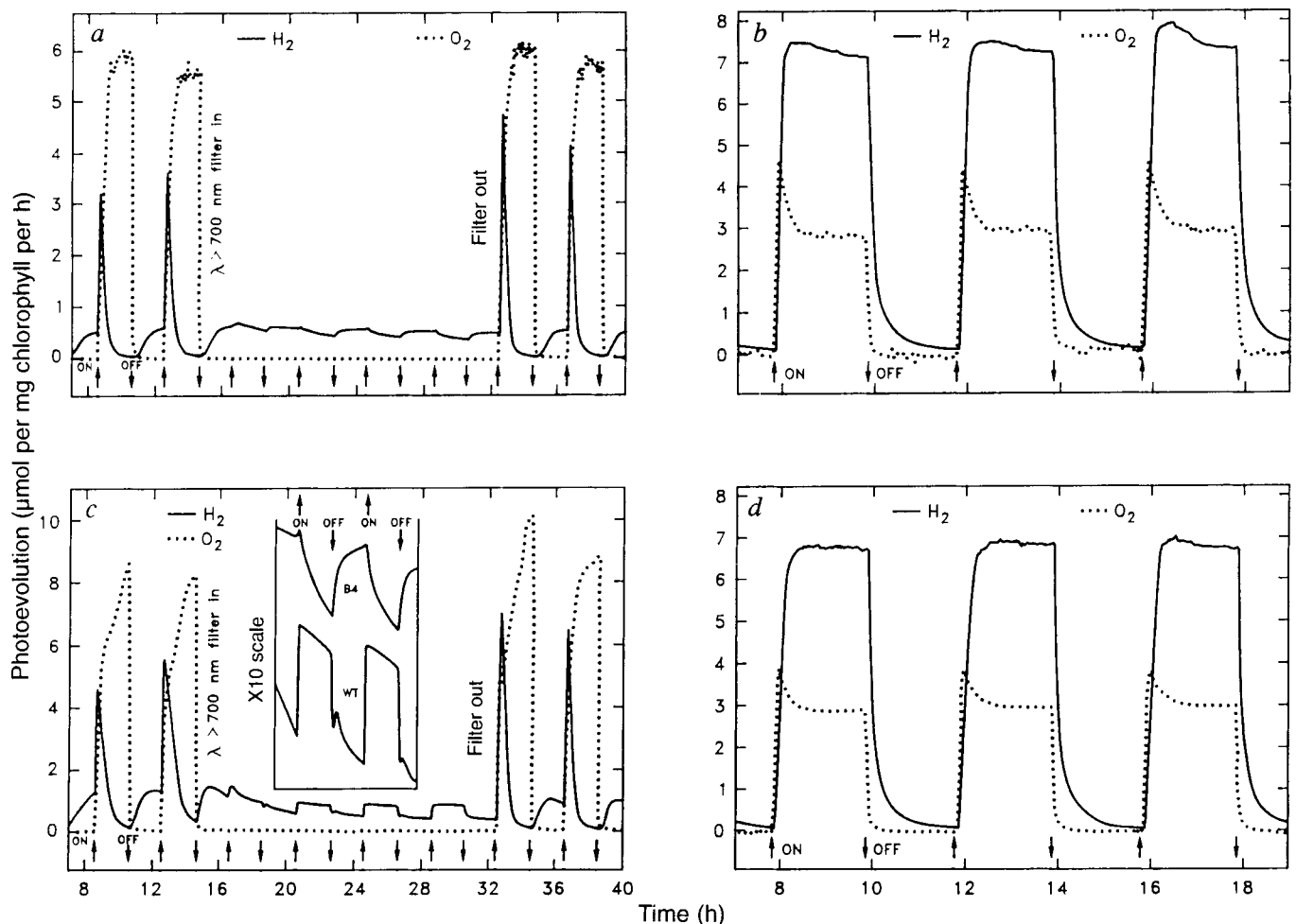


FIG. 1 Simultaneous photoevolution of molecular hydrogen and oxygen by *Chlamydomonas* mutant B4 and wild-type *Chlamydomonas* 137c with non-saturating light. a, Mutant B4, 363 p.p.m. CO₂ in helium; b, mutant B4, helium; c, wild-type, 363 p.p.m. CO₂ in helium; d, wild-type, helium. As indicated in the insert of c, irradiation with light of $\lambda > 700$ nm generated PS-I-dependent hydrogen evolution in strain

137c, whereas no such hydrogen evolution was observed in mutant B4. In a and c, the peak and decline of hydrogen evolution while the light is still on represents the real-time activation of the Calvin–Benson cycle after a period of darkness. The chlorophyll concentration for the experiments shown in a and c was 15 $\mu\text{g ml}^{-1}$; for b and d it was 3 $\mu\text{g ml}^{-1}$. In each case, the volume of illuminated algal suspension was 25 ml.

monotonic decrease in the rate of hydrogen production from its peak value represents the real-time activation of the CO_2 reduction cycle. When the light is turned off, thermally activated hydrogen evolution is observed. This pattern repeats itself when the light is turned on again. The source of thermally activated hydrogen is reduced carbon, identified⁸ as primarily starch that has been stored during previous photosynthesis. *C. reinhardtii* grows quite well in an anaerobic atmosphere consisting of CO_2 and helium⁹. In a pure-helium atmosphere, illustrated by the data in Fig. 1b, d, the hydrogen time profiles are qualitatively different. They climb monotonically to an approximate steady-state production rate with the light turned on and return to zero when it is turned off. Because CO_2 is not present in this atmosphere, the CO_2 -reduction cycle is not activated and does not serve as a sink for reducing equivalents. Oxygen-evolving activity is less in helium, presumably due to the absence of CO_2 and the requirement of PS II for bicarbonate¹⁰. Irradiation with $\lambda > 700$ nm light (which supports PS I but not PS II) in the wild-type alga (Fig. 1c insert) results in a low, but significant, rate of hydrogen photoevolution. The source of reducing equivalents for this PS I-dependent hydrogen production is the pool of

reduced electron carriers in the electron transport chain linking the two photosystems. Under strictly anaerobic conditions, this pool is replenished by the endogenous reductants described above. B4 irradiated at $\lambda > 700$ nm (Fig. 1a, also shown in Fig. 1c insert on a $\times 10$ expanded scale), shows, as expected, no PS I hydrogen-production component. Instead, thermally activated hydrogen evolution was suppressed. These control data, in conjunction with the controls described above, indicated the absence of PS I before, during and after the experiments.

Perhaps the most remarkable aspect of the data for mutant B4 for CO_2 in helium (Fig. 1a) is that the Calvin-Benson cycle is evidently activated in this system and serves as an electron sink. This is positively demonstrated by the data of Fig. 3 in which the photoassimilation of CO_2 by both mutant and wild type is illustrated. It therefore follows that mutant B4 can perform complete oxygenic photosynthesis using a single PS II reaction. As shown in Fig. 3a, at saturating-high intensity (500 W m^{-2} photosynthetically active radiation), CO_2 photoassimilation is stable for both mutant B4 and wild-type *Chlamydomonas* in an anaerobic atmosphere. However, as shown in Fig. 3b, only the wild-type alga, containing both PS I and PS II, is stable in air. Within 24 hours, mutant B4 lost 90% of its photosynthetic activity. Figure 2 illustrates the post-irradiated algae, cultures of wild-type and mutant B4 in air, with the latter completely bleached. Thus we concluded that PS I is required for stable oxygenic photosynthesis in aerobic environments (the normal condition in nature) but not under anaerobic conditions. As the oxidant generated in PS II must be more oxidizing than $+0.8 \text{ V}$, it follows that a single PS II reaction centre has the thermodynamic potential to perform complete photosynthesis, as does the energy content of a single photon absorbed by PS II. Thus, although our findings are novel and unexpected, they do not violate known thermodynamic requirements of the photosynthetic process.

A number of observations in the literature have suggested that PS II alone could reduce ferredoxin and thus NADP^+ . The photoreduction of NADP^+ has been observed¹¹ in *Chlamydomonas* ACC-1, a mutant lacking PS I. However, oxygen evolution was not demonstrated, and this was interpreted as a non-physiological reaction, an alternative electron pathway occurring only at high light intensities. The transient evolution of both hydrogen and oxygen from a mutant of *Chlamydomonas* that was deficient in PS I has been observed¹² by a polarographic technique. Arnon and Barber¹³ reported the photoreduction of NADP^+ by isolated reaction centres of PS II with 1,5-diphenylcarbazide as the electron source. The reaction centres of that work were not functional oxygen evolution preparations; the key emphasis of the Arnon and Barber paper was the requirement for plastocyanin. More recently, Allakhverdiev and Klimov¹⁴ demonstrated the requirement of manganese for anaerobic photoreduction of NADP^+ by PS II of higher plants. The primary electron acceptor of PS II is pheophytin, which has an E_m of -0.61 V (refs 15–17).

Irrespective of specific molecular mechanisms, our data clearly illustrate a new type of photosynthesis being performed by the PS II light reaction exclusively. These experiments also have implications for the evolution and origin of water-splitting photosynthesis. Anoxic photosynthetic bacteria exhibit two different mechanisms characterized by either pheophytin-quinone or Fe-S type reaction centres. Blankenship¹⁸ proposed that "some sort of genetic fusion event took place between two bacteria", producing a chimaeric organism. Subsequently, the two systems were linked, and the oxygen evolution system was added. However, no examples of the intermediate stages have been observed. Our data support an alternative hypothesis: a bacterium containing the pheophytin-quinone reaction centre developed into the oxygen-evolving PS II type of photosynthesis, with PS I being added later for survival when molecular oxygen became a major component of the Earth's atmosphere. Data that support this theory are presented in Fig. 3a, b. The photo-

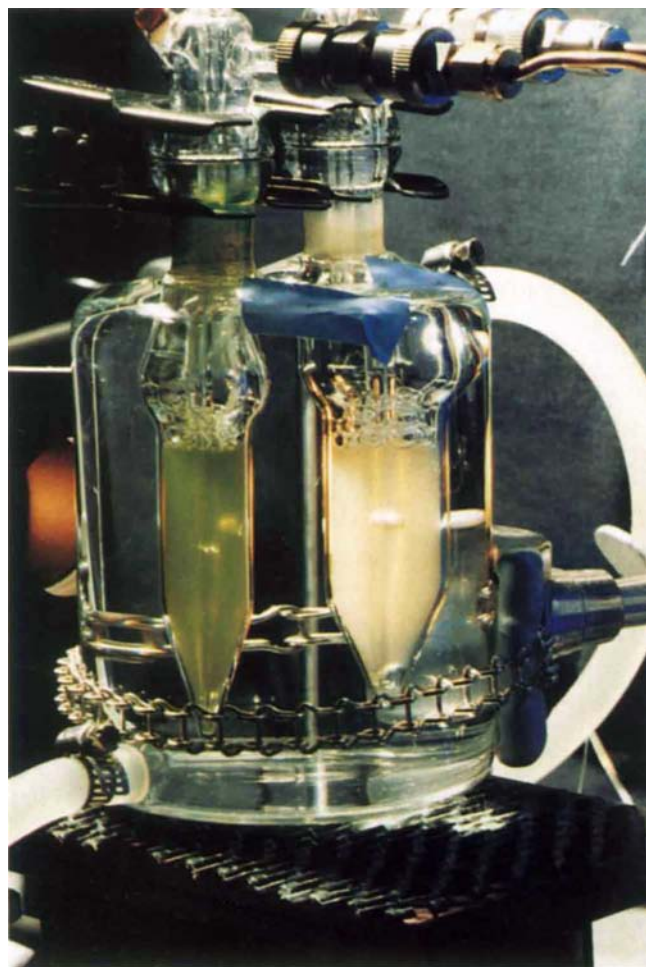


FIG. 2 Photograph of post-irradiated algae in air with 500 W m^{-2} photosynthetically active light corresponding to the data of Fig. 3b. Left, wild-type *Chlamydomonas* 137c; right, *Chlamydomonas* mutant B4. Although the initial chlorophyll concentration of each sample was the same at the start of the experiment, the final concentration of the wild-type alga increased after irradiation while B4's chlorophyll diminished. The algae were illuminated for 23 h at 20°C . The algal chamber on the right containing the bleached sample of B4 appears larger in diameter owing to the lens effect of the water jacket surrounding both identical chambers.

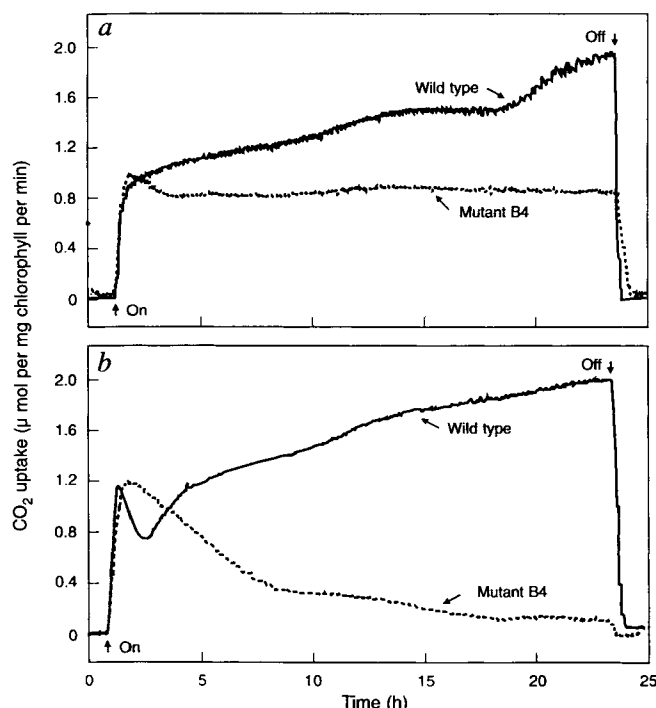


FIG. 3 A comparative study of photosynthesis between mutant B4 and wild-type 137c in anaerobic and aerobic atmospheres: a, 363 p.p.m. CO_2 in helium; b, 363 p.p.m. CO_2 in air. In each case, the chlorophyll concentration was $15 \mu\text{g ml}^{-1}$ and the volume of illuminated algal suspension was 13 ml. The loss of photosynthetic activity by mutant B4 indicated in b was accompanied by a bleaching of its chlorophyll. See Fig. 2 for a visual comparison. Continuous measurement of CO_2 was achieved with an in-line infrared spectrophotometer. The CO_2 detector was calibrated with an electronic gas blender with primary standard gas mixtures of CO_2 in helium. Because measurements on both mutant and wild-type were performed simultaneously under identical optical and geometrical conditions, it is clear from these data and those of Fig. 1 that the relative quantum yields for the two systems are comparable.

production of superoxide radicals by PS II membrane fragments has recently been demonstrated¹⁹. This may be related to our unsuccessful earlier attempts at observing hydrogen photoproduction by platinized PS II preparations (ref. 20, and E.G. and M. Seibert, unpublished observations).

From a practical point of view, a single-light reaction implies that the maximum thermodynamic conversion efficiency of light energy into chemical energy, as represented by the Gibbs free energy of molecular hydrogen, can be potentially doubled, from about 10% to 20%, because a single photon rather than two is now required to span the potential difference between water oxidation/oxygen evolution and proton reduction/hydrogen evolution. This is obviously not true of the data in Fig. 1 where, under identical conditions, B4 and wild-type *Chlamydomonas* have similar quantum efficiencies. But further studies on algal mutants may achieve the desired result of increased light conversion efficiency. □

Received 16 January; accepted 6 June 1995.

1. Roitgrund, C. & Mets, L. *Curr. Genet.* **17**, 147–153 (1990).
2. Bennoun, P. & Chua, N.-H. in *Genetics and Biogenesis of Chloroplasts and Mitochondria* (eds Böcher, T., Neupert, W., Sebald, W. & Werner, S.) 33–39 (North-Holland, Amsterdam, 1976).
3. Choquet, Y. et al. *Cell* **52**, 903–913 (1988).
4. Greenbaum, E. *Photobiophys. Photobiophys.* **8**, 323–332 (1984).
5. Latimer, W. M. in *The Oxidation States of the Elements and Their Potentials in Aqueous Solutions*, 2nd edn, 29–50 (Prentice-Hall, New York, 1952).
6. Bouges-Bouquet, B. *Biochim. biophys. Acta* **594**, 85–103 (1980).
7. Graves, D. A., Tevault, C. V. & Greenbaum, E. *Photochem. Photobiol.* **50**, 571–576 (1989).

8. Klein, U. & Betz, A. *Pl. Physiol.* **61**, 953–956 (1978).
9. Graves, D. A., Spradlin, G. M. & Greenbaum, E. *Photochem. Photobiol.* **51**, 585–590 (1990).
10. Stemler, A. & Govindjee *Planta* **52**, 119–123 (1973).
11. Klimov, V. V., Allakhverdiev, S. I. & Ladygin, V. G. *Photosyn. Res.* **10**, 355–361 (1986).
12. Boychenko, V. A., Allakhverdiev, S. I., Ladygin, V. G. & Klimov, V. V. *Dokl. Akad. Nauk SSSR* 955–998 (1986).
13. Arnon, D. I. & Barber, J. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5930–5934 (1990).
14. Allakhverdiev, S. I. & Klimov, V. V. *Z. Naturf.* **47c**, 57–62 (1992).
15. Klimov, V. V., Klevanik, A. V., Shuvalov, V. A. & Krasnovsky, A. A. *FEBS Lett.* **82**, 183–186 (1977).
16. Klimov, V. V. & Krasnovsky, A. A. *Photosynthetica* **15**, 592–609 (1981).
17. Klimov, V. V., Allakhverdiev, S. I., Dementev, S. & Krasnovsky, A. A. *Dokl. Akad. Nauk SSSR* **249**, 227–230 (1979).
18. Blankenship, R. E. *Photosyn. Res.* **33**, 91–111 (1992).
19. Ananyev, G., Renger, G., Wacker, U. & Klimov, V. *Photosyn. Res.* **41**, 327–338 (1994).
20. Woodward, J., Lewis, B., Miracle, G. & Greenbaum, E. *Appl. Biochem. Biotechnol.* **20/21**, 259–265 (1989).

ACKNOWLEDGEMENTS. This paper is dedicated to Professor D. I. Arnon (1910–1995). We thank J. R. Benemann, A. San Pietro, M. Seibert and J. Barber for reading the manuscript and making comments, suggestions and criticisms. We also thank D. Zhang and M. Thumauer for independently confirming, by EPR spectroscopy, the lack of detectable PS I in mutant B4; T. Owens for demonstrating the lack of P700 in B4 with optical absorbance spectroscopy; and A. Pascual for technical support. Mutant B4 has been deposited with the Duke University *Chlamydomonas* Genetics Center as Stock No. CC2933. This work was supported by the U.S. Department of Energy, the Pittsburgh Energy Technology Center, and the NSF. Oak Ridge National Laboratory is managed by Lockheed Martin Energy Research Corp., for the U.S. Department of Energy.

A model of protein synthesis based on cryo-electron microscopy of the *E. coli* ribosome

Joachim Frank^{*†}, Jun Zhu^{*†}, Pawel Penczek^{*}, Yanhong Li^{*†}, Suman Srivastava^{*}, Adriana Verschoor^{*}, Michael Rademacher^{*†}, Robert Grassucci^{*}, Ramani K. Lata^{*} & Rajendra K. Agrawal^{*}

^{*} Wadsworth Center, New York State Department of Health, Empire State Plaza, PO Box 509, Albany, New York 12201-0509, USA
[†] Departments of Biomedical Sciences and [‡] Computer Science, State University of New York at Albany, 1400 Washington Avenue, Albany, New York 12222, USA

THE ribosome is formed by assembly of proteins and nucleic acids, and synthesizes proteins according to genetic instructions in all organisms. Many of the biochemical steps of this fundamental process are known, but a detailed understanding requires a well-defined structural model of the ribosome. Electron microscopy combined with image reconstruction of two-dimensional crystals^{1–3} or single ribosomes⁴ has been the most promising technique, but the resolution of the resulting models has been insufficient. Here we report a 25-Å reconstruction of the ribosome from *Escherichia coli*, obtained by combining 4,300 projections of ice-embedded single particles. Our new reconstruction reveals a channel in the small ribosomal subunit and a bifurcating tunnel in the large subunit which may constitute pathways for the incoming message and the nascent polypeptide chain, respectively. Based on these new findings, a three-dimensional model of the basic framework of protein synthesis is presented.

An energy-filtering electron microscope was used to visualize ribosomes prepared as frozen hydrated samples⁵ (Fig. 1). Our reconstruction (Fig. 2a–c) shows the two ribosomal subunits as well defined separate masses of density that are joined by five bridges across the subunit-subunit interface. Morphological features established for the two subunits are easily recognized: the head (h) and platform (p) for the 30S subunit, and the L1 arm (L1), central protuberance (CP), stalk (St), and interface canyon (IC) (ref. 6) for the 50S subunit. A new spur-like feature (sp) appears at the base of the small subunit⁷. The subunits enclose a space ('intersubunit space') that accommodates mes-

FIG. 1 Images of ice-embedded ribosomes and orientations of the 4,300 selected particles. *a*, Portion of a micrograph showing ribosomes in random orientations (scale bar, 1,000 Å). *b*, Orientations of the particles. Each particle is represented by a point in a (θ, ϕ) angular coordinate system. *c*, Averages of clusters found by classification.

METHODS. Inelastic scattering of the electrons in the ice causes a loss of contrast and makes it difficult to correct the electron microscope's contrast transfer function. The use of an energy-filtering electron microscope eliminates these problems entirely²¹. Micrographs (20) of untilted specimen grids were recorded under low-dose conditions, using different defocus settings (Δz , 2.0 μm and 2.5 μm) and an energy window of 14 eV. Particles (2,043 and 2,254, respectively) were picked and subjected to reference-free alignment²² and classification²³. The orientation of each projection was determined individually by matching it⁷ to computed projections of the previous 29-Å model⁷ that was obtained without energy filtering.

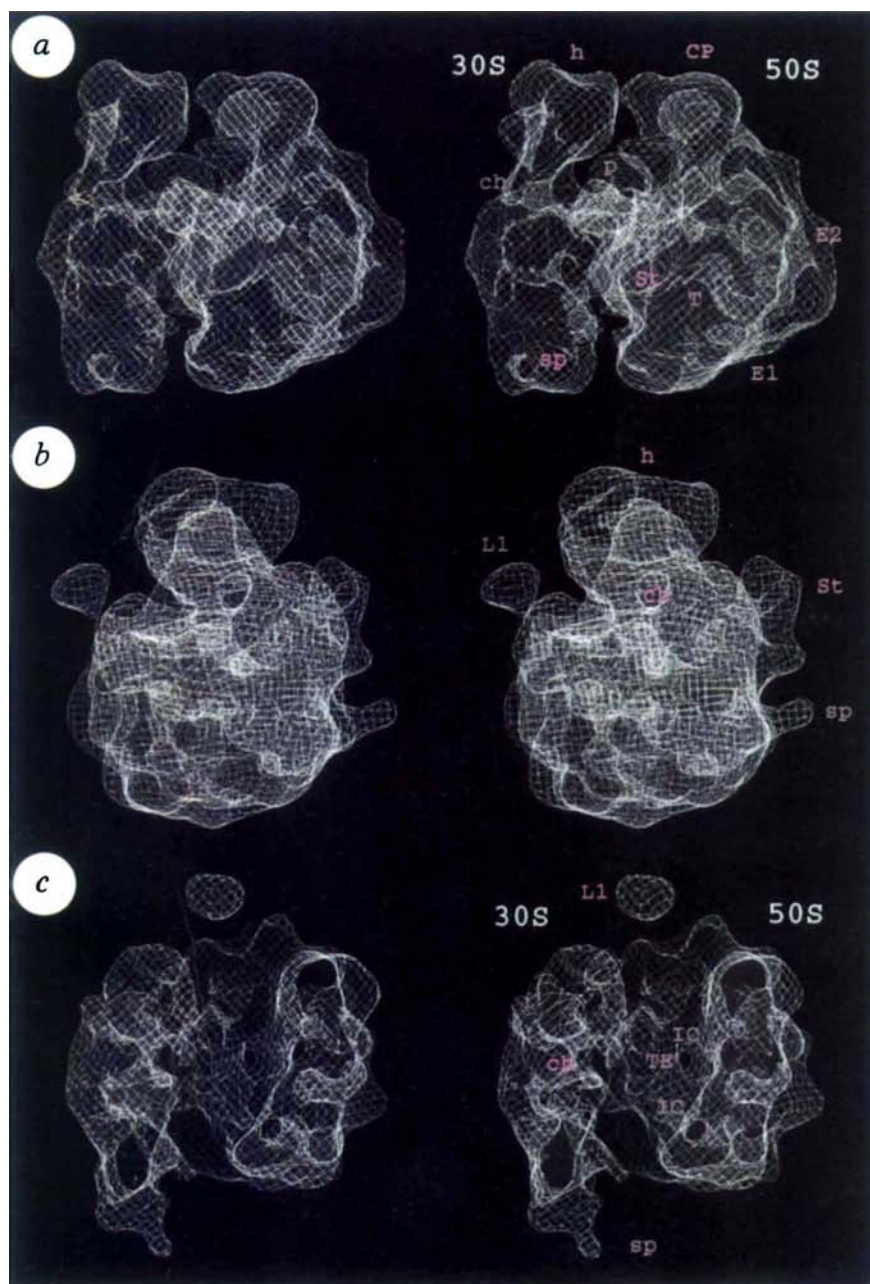
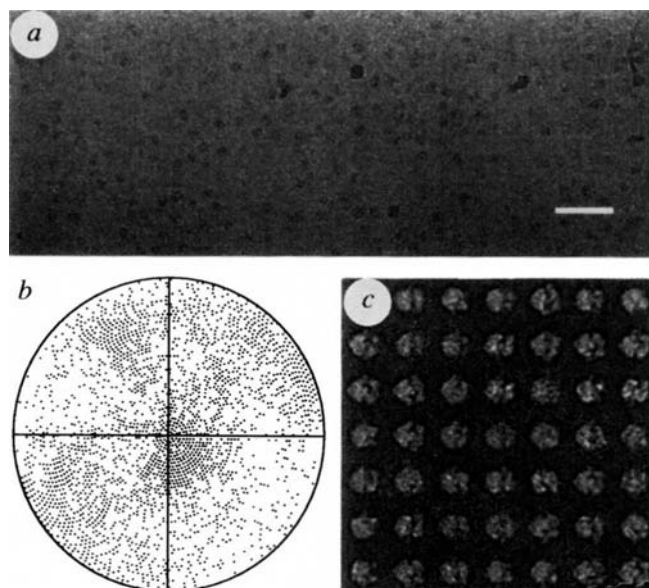
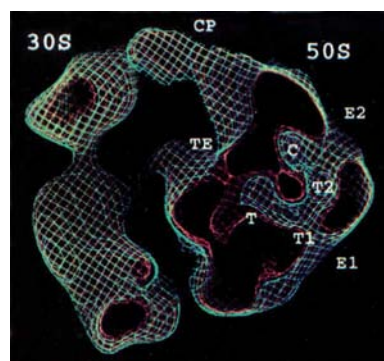


FIG. 2 Three-dimensional map of the *E. coli* ribosome presented from different viewing angles. Lower-case letters refer to features of the 30S subunit: h, head; p, platform; ch, channel; sp, spur. Upper-case letters refer to features of the 50S subunit: L1, L1 arm; CP, central protuberance; St, stalk; IC, interface canyon; T, tunnel; TE, tunnel entry; E1 and E2, exit sites of tunnel after branching. *a*, Side view showing part of the tunnel through the 50S subunit. The tunnel connects with the bottom of the interface canyon at higher contouring levels (see *c* and Fig. 3). *b*, View down the channel through the neck of the 30S subunit (generated from *a* by rotation by 90° around vertical axis, then by -15° around the horizontal axis). *c*, View down the tunnel through the 50S subunit (generated from *a* by rotation by 75° around horizontal axis and use of cutting plane to remove part of the ribosome). A higher contour level was used to visualize tunnel entry. **METHODS.** The reconstruction was obtained by the simultaneous iterative reconstruction technique (SIRT)^{22,24}. To verify that the solution is stable and unique, we wobbled the angles of the projections randomly by 15°, and refined against the resulting 45-Å low-resolution reconstruction using the 3D projection alignment procedure⁷. Two such refinements produced virtually identical results. In addition, an independent reconstruction from cluster averages using the method of common lines and a new minimization technique further confirmed the results (P.P., J.Z. and J.F., in preparation). The two sets of projections were combined in a new iterative 3D reconstruction procedure, in which the different contrast transfer functions (CTFs) of the electron microscope were explicitly taken into account (J.Z., P.P., R. Schröder and J.F., in preparation). The most important effect of the CTF correction is that it restores the correct weight of the low spatial frequencies, which carry information on the relative magnitudes of densities that are separated by large distances. Uncorrected reconstructions show several tunnels, but some of these are found to possess only slightly reduced density when the volume is CTF-corrected. Similarly, the identification of RNA-rich regions on the basis of their higher scattering density can be made with greater certainty on such a corrected volume.

FIG. 3 Cross-section of the 3D map showing the bifurcating tunnel in the 50S subunit. Two different contour levels were used. For the normal contouring level (green), the primary tunnel (T) ends upwards in a cul-de-sac, probably because the upper segment has a diameter substantially smaller than the resolution. If present, such a thin tunnel segment will still be recognized as a line of low density. Indeed, the high contouring level (red) shows a connection of the tunnel with the bottom of the interface canyon (TE). (Other expansions of the tunnels, visible in green, can be explained by the removal of protein.) CP, central protuberance; T1 and T2, lower and upper tunnel branches, respectively; C, side compartment; E1 and E2, tunnel exit sites.



senger RNA, transfer RNAs and associated factors during protein synthesis.

The new three-dimensional density map contains several new features that allow tentative identification of the pathways of the incoming mRNA and the exiting nascent polypeptide, in turn suggesting the locations of the anticodon-binding region and peptidyltransferase centre: (1) the 'neck' portion of the small subunit is penetrated by a channel (ch in Fig. 2a, b) leading from the cytoplasmic side into the intersubunit space; (2) the cup-shaped platform of the small subunit (p in Fig. 2a) forms a well defined cleft with the main body, into which the channel leads, and (3) a tunnel (T), whose upper portion is seen at elevated contouring levels, starts at the bottom of the interface canyon (tunnel entry, TE, in Figs 2c, and 3) and bifurcates into two tunnels (T1, T2 in Fig. 3), which exit at different points (E1, E2) on the back of the 50S subunit. In addition, a compartment (C) is recognized (Fig. 3). A single, low-resolution tunnel through the 50S subunit was previously observed^{1,2,8}.

Our interpretation (Fig. 4) is based on the locations of the new features relative to sites previously mapped, and on their morphologies. The incoming mRNA threads through the channel in the small-subunit neck, makes a U-turn⁹ and exits by a pathway circumscribed by the platform rim. Threading of the mRNA through the neck has been postulated¹⁰ as a means of unfolding the mRNA into a single-stranded form. In fact, the channel is located in a region of the subunit that is thought to be traversed by a single-stranded segment of the 16S ribosomal RNA^{11,12}. The suggested locations of entry and exit of the mRNA restrict the possible sites of binding of the aminoacyl and peptidyl tRNAs (A and P sites) to a region 20 Å wide in the cleft, to which a 16S ribosomal RNA region implicated in mRNA binding has been mapped¹³.

In agreement with previous observations and conjectures^{1-3,8,14,15}, we interpret the large-subunit tunnel as the conduit through which the nascent polypeptide exits the ribosome. However, our results suggest the existence of not one but two

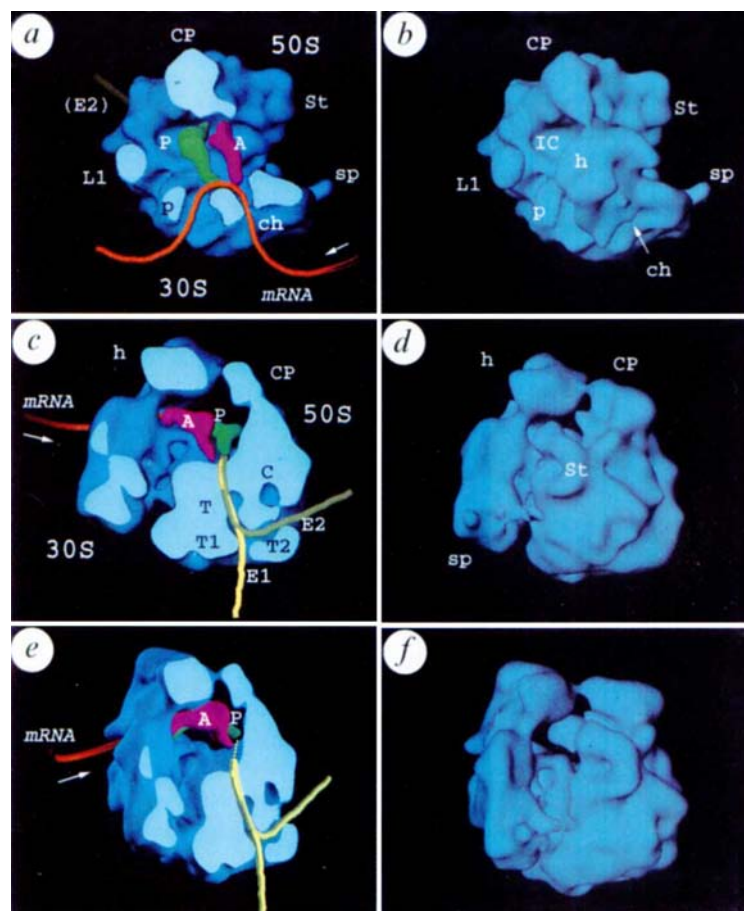


FIG. 4 Model (3D) of the basic framework of protein synthesis using a surface representation of the 25-Å reconstruction. The 12.5 Å map of tRNA-tRNA complex¹⁹ (A site, pink; P site, green) was placed such that the anticodon ends face the cleft region immediately next to the channel entry. Aminoacyl ends were placed at two extreme positions related by rotation of the complex around mRNA axis within the range allowed by space exclusion. The loop-shaped mRNA pathway entering the channel from cytosol and encircling the back portion of neck is shown in orange. Two possible pathways of exit of the polypeptide chain are shown in olive green and yellow. a, Plane that sections the channel (ch) and the cleft that it leads into. L1, L1 arm; CP, central protuberance; St, stalk; p, platform; sp, spur. The tRNA complex (A and P) has been placed such that the anticodon ends are in the cleft, and aminoacyl ends at the tunnel entry as shown in c. Olive strand marked (E2) is a possible path of polypeptide chain exiting through the opening E2 that is marked in c. b, Corresponding top view of whole ribosome. IC, interface canyon; h, head of small subunit. c, Plane that sections the bifurcating tunnel (T, T1, T2). C, compartment contiguous with T2; E1 and E2, alternative exit sites. The tRNA complex is placed with aminoacyl ends at the lowest possible position. d, Corresponding side view of whole ribosome. e, Sectioning plane as in c, but the tRNA complex is placed with aminoacyl ends at the highest possible position limited by contact to the 'roof' formed by the small subunit head. The model has been rotated around the vertical axis, allowing a glancing view along the canyon wall of the large subunit. The broken line marks the range of possible positions of the peptidyltransferase centre along that wall. f, Corresponding side view of the whole ribosome.

pathways, created by a bifurcation at the bottom of the primary tunnel T into a short lower tunnel, T1, and a long upper tunnel, T2. The short branch T1 (Fig. 3) emerges at the large-subunit base, within the 'exit domain' delineated by immunoelectron microscopy¹⁴, whereas the long branch T2 emerges towards the middle of the subunit back, approximately 65 Å above the first.

The main pathway (T, T1 in Fig. 3) has a linear length of 85 Å, roughly equivalent to a 40-residue α -helical segment of polypeptide, the span found to be protected from proteases¹⁶. The exit of this tunnel, E1, lies in a domain of the large subunit believed to be tightly associated with the cytoplasmic membrane¹⁷, forming part of the secretory pathway. The length of the branch T2, from the point of bifurcation to its exit, E2, is 50 Å. Towards E2 it becomes contiguous with a globular compartment, C (Figs 3 and 4b, c). For a membrane-bound ribosome, E2 would lead into the cytoplasm. However, the way in which two alternative pathways could be used in translocation (see, for example, ref. 18) is unknown.

The tunnel entry point (TE in Fig. 3) at the bottom of the interface canyon implies a location of the peptidyltransferase centre somewhere above this point. To restrict the range of possible locations of the peptidyltransferase centre we must dock A- and P-site tRNAs to the putative anticodon-binding region in the cleft, and further consider the spatial constraints imposed by the intersubunit space and the steric interactions between mRNA and A- and P-site tRNAs.

We have used the atomic model of the A-tRNA-P-tRNA complex¹⁹ in S configuration²⁰ to produce a low-resolution image, as would be obtained by an electron microscope with 12.5 Å resolution. When the complex is docked onto the putative codon-binding site (Fig. 4a), it has freedom to rotate within the intersubunit space by 30° around the axis defined by the two codons. Two extreme positions are shown (Fig. 4c, e). In the first, the aminoacyl end of the P-site tRNA is proximal to the tunnel entry; in the other extreme, it is about 25 Å above this point. Either position would leave ample space towards the back of the A- and P-tRNA complex for an additional E-site tRNA. Detailed three-dimensional mapping experiments are needed to confirm our hypothesis and to allow refinement of our basic model of protein synthesis. □

Sequestration of the membrane-targeting myristoyl group of recoverin in the calcium-free state

Toshiyuki Tanaka*, James B. Ames†, Timothy S. Harvey*†, Lubert Stryer†§ & Mitsuhiro Ikura*§

* Division of Molecular and Structural Biology, Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, 500 Sherbourne Street, Toronto, Ontario M4X 1K9, Canada

† Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA

RECOVERIN, a retinal calcium-binding protein of relative molecular mass (M_r) 23K, participates in the recovery phase of visual excitation and in adaptation to background light¹⁻³. The Ca^{2+} -bound form of recoverin prolongs the photoresponse⁴, probably by blocking phosphorylation of photoexcited rhodopsin⁵. Retinal recoverin contains a covalently attached myristoyl group or related acyl group at its amino terminus⁶ and two Ca^{2+} -binding sites⁷. Ca^{2+} binding to myristoylated, but not unmyristoylated, recoverin induces its translocation to bilayer membranes, indicating that the myristoyl group is essential to the read-out of calcium signals (calcium-myristoyl switch)^{8,9}. Here we present the solution structure of Ca^{2+} -free, myristoylated recombinant recoverin obtained by heteronuclear multidimensional NMR spectroscopy. The myristoyl group is sequestered in a deep hydrophobic pocket formed by many aromatic and other hydrophobic residues from five flanking helices.

A best-fit superposition of the 24 NMR-derived structures is shown in Fig. 1a. The structure contains 11 α -helices (denoted as A-K in Fig. 1b) and two pairs of short antiparallel β -sheets. Four pairs of helices constitute four EF-hand type helix-loop-helix structural motifs (EF-1 and EF-2 in the N-terminal domain, and EF-3 and EF-4 in the carboxy-terminal domain). A sharply bent linker (residues 92 to 97) between the two domains allows the four EF-hands to form a compact linear array that contrasts with the dumb-bell arrangement seen in calmodulin¹⁰ and troponin C (ref. 11).

The myristoyl group attached to the N-terminal glycine adopts an extended conformation (Fig. 1b) that projects into a deep hydrophobic pocket located in the N-terminal domain. This linear conformation contrasts with the curved shapes of the myristoyl group found in the crystal structure of cyclic AMP-dependent protein kinase¹² and in that of the Mahoney strain of type 1 poliovirus¹³. Hydrophobic amino acid residues that are spatially close to the myristoyl group and contribute to the binding pocket are depicted in Fig. 2a. These key residues are provided by five helices (Fig. 2b). Helices B, C, E and F surround the myristoyl group laterally and provide a skeletal frame for the binding pocket. Helix A, which makes contact with one face of the myristoyl group, serves as a lid on top of the other four helices. The N-terminal amino acid residues Gly 2 and Asn 3 form a tight hairpin turn adjacent to the N-terminal helix. This hairpin turn, which resembles a hook, positions the myristoyl group inward and antiparallel to the N-terminal helix. A similar hairpin-like but longer turn was observed in the crystal structure of myristoylated cAMP-dependent protein kinase¹², which also has a long helix adjacent to the myristate in an antiparallel fashion. Helix D, the only helix in the N-terminal domain that

Received 12 May; accepted 23 June 1995.

- Yonath, A., Leonard, K. R. & Wittmann, H. G. *Science* **236**, 813-816 (1987).
- Eisenstein, M. et al. *Biochimie* **73**, 879-886 (1991).
- Yonath, A. & Wittmann, H. G. *Trends biochem. Sci.* **14**, 329-335 (1989).
- Frank, J., Penczek, P., Grassucci, R. & Srivastava, S. J. *Cell Biol.* **115**, 597-605 (1991).
- Lepault, J., Booy, F. P. & Dubochet, J. *J. Microsc.* **129**, 89-102 (1983).
- Radermacher, M., Wagenknecht, T., Verschoor, A. & Frank, J. *EMBO J.* **6**, 1107-1114 (1987).
- Penczek, P., Grassucci, R. & Frank, J. *Ultramicroscopy* **53**, 251-270 (1994).
- Frank, J., Verschoor, A., Radermacher, M. & Wagenknecht, T. in *Ribosomes* (eds Hill, W. E. et al.) 107-113 (American Society for Microbiology, Washington DC, 1990).
- Montesano-Roditis, L. & Giltz, D. G. *J. Biol. Chem.* **269**, 6458-6470 (1994).
- Lim, V. et al. *Nucleic Acids Res.* **20**, 2627-2637 (1992).
- Shatsky, I., Bakin, A. V., Bogdanov, A. A. & Vasiliev, V. D. *Biochimie* **73**, 937-945 (1991).
- Malhotra, A. & Harvey, S. C. *J. molec. Biol.* **240**, 308-340 (1994).
- Oakes, M. I. & Lake, J. A. *J. molec. Biol.* **211**, 897-906 (1990).
- Bernabeu, C. & Lake, J. A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3111-3115 (1982).
- Avila-Sakar, A. J. et al. *J. molec. Biol.* **239**, 689-697 (1994).
- Biobel, G. & Sabatini, D. D. *J. Cell Biol.* **48**, 130-145 (1970).
- Crowley, K. S., Reinhart, G. D. & Johnson, A. E. *Cell* **73**, 1101-1115 (1993).
- Kudlicki, W. et al. *J. Biol. Chem.* **270**, 10650-10657 (1995).
- Easterwood, T. R., Major, F., Malhotra, A. & Harvey, S. C. *Nucleic Acids Res.* **22**, 3779-3786 (1994).
- Sundaralingam, M., Brenman, T., Yathindra, N. & Ichikawa, T. in *Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions* (eds Sundaralingam, M. & Rao, S. T.) 101-115 (University Park, Baltimore, 1975).
- Schröder, R. R., Hofmann, W. & Menetret, J.-F. *J. struct. Biol.* **105**, 28-34 (1990).
- Penczek, P., Radermacher, M. & Frank, J. *Ultramicroscopy* **40**, 33-53 (1992).
- Frank, J. Q. *Rev. Biophys.* **23**, 281-329 (1990).
- Gilbert, P. J. *theor. Biol.* **36**, 105-117 (1972).

ACKNOWLEDGEMENTS J.F. and J.Z. thank K. Holmes and R. Schröder at the Max Planck Institute for Medical Research, Biophysics Department for hospitality during J.F.'s sabbatical visit supported by the Alexander-von-Humboldt-Stiftung and NIH, and for the use of the Zeiss EM 912 electron microscope. We acknowledge use of the Wadsworth Center's Research Computing Core. We thank R. Brimacombe for discussion of the tRNA fitting; S. Harvey for providing the atomic coordinates of the tRNA-tRNA model; A. Leith for assistance with the modelling; and T. Wagenknecht for discussion. This work was supported by grants from NIH and NSF to J.F.

† Present address: Molecular Structure Group, Amgen Inc., 1840 DeHavilland Drive, Thousand Oaks, California 91320, USA.

§ To whom correspondence should be addressed.

has no contact with the myristoyl group, seems to be the most flexible of all the EF-hand helices¹⁴.

A number of recently discovered recoverin homologues are all acylated at their N termini^{15–18}. A moderate to high degree of sequence identity is found between recoverin and homologues such as S-modulin (83%), visinin (59%), hippocalcin (49%), neurocalcin (49%), neural visinin-like protein (44%) and frequenin (39%). Numerous hydrophobic residues are conserved among the recoverin subfamily members (Fig. 3). In particular, the aromatic residues Trp 31, Phe 49, Tyr 53, Phe 56, Phe 57, Phe 83 and Trp 104, which contribute to the myristoyl-binding pocket (Fig. 2a), remain invariant. This strongly suggests that the structure of the binding pocket is similar in all members of the subfamily. The homologues also have several basic residues at the N terminus, which may help anchor the proteins to membranes by electrostatic interactions with acidic phospholipids¹⁹. By contrast, significant differences in amino-acid sequence are found at the C terminus (residues 191 to 202 in recoverin). Recoverin and S-modulin have a longer C-terminal tail than the other subfamily members. This region is rich in positively charged residues, which may also interact electrostatically with phospholipid membranes. The members of the subfamily are found exclusively in the nervous system. Several have been shown to contain a calcium-myristoyl switch^{8,9,16,17}. They prob-

ably participate in membrane-associated signal transduction processes by coupling G-protein receptors to Ca^{2+} cascades^{1,2}.

It is illuminating to compare the solution structure of Ca^{2+} -free, myristoylated recoverin with the crystal structure of single- Ca^{2+} -bound unmyristoylated recoverin²⁰. Their C-terminal domains, which contain EF-3 and EF-4, are relatively similar (backbone root-mean-square deviation (r.m.s.d.) of 2.2 ± 0.1 Å), whereas their N-terminal domains, which contain EF-1 and EF-2, are considerably different (backbone r.m.s.d. of 6.9 ± 0.1 Å). First, the helices of EF-1 are nearly antiparallel ($169 \pm 2^\circ$) in contrast with the nearly perpendicular orientation (108°) observed in the crystal structure. Substantially smaller differences in the inter-helix angle are observed for EF-2, EF-3 and EF-4 ($136 \pm 3^\circ$, $117 \pm 2^\circ$ and $98 \pm 3^\circ$ in the Ca^{2+} -free solution structure, and 129° , 101° and 97° in the single- Ca^{2+} -bound crystal structure, respectively). Second, the eight-residue loop (residues 17–24) between EF-1 and the N-terminal helix A allows different positions of helix A with respect to helix B. Third, helix A (residues 4 to 16) is stabilized by the myristoyl group in the Ca^{2+} -free structure, whereas the first seven residues (residues 2–8) are unstructured in the crystal structure. Furthermore, the relative orientation of the N-terminal and C-terminal domains is strikingly different between the two forms of recoverin (backbone r.m.s.d. of 9.6 ± 0.2 Å for the entire molecule). In essence,



FIG. 1 a, Stereoview of a best-fit superposition of the backbone (N, Ca and C) atoms of the 24 NMR-derived structures of Ca^{2+} -free, myristoylated recoverin, generated using Insight II (Biosym, San Diego). The main-chain atoms of residues 1–94 of the 24 structures are superimposed against the energy-minimized average structure of residues 1–189 using residues 4–16, 25–38, 43–56, 66–73 and 80–91 (root-mean-square deviation (r.m.s.d.) of 0.44 ± 0.04 Å for backbone atoms and 1.07 ± 0.06 Å for all heavy atoms). The main chains of residues 95–189 of the same set of structures are superimposed against the average structure using 98–109, 116–132, 148–159, 166–178 and 180–186 (the r.m.s.d. values are 0.42 ± 0.06 Å for backbone atoms and 0.99 ± 0.07 Å for all heavy atoms). The r.m.s.d. values for the backbone and all heavy atoms of the entire protein are 0.61 ± 0.14 Å and 1.15 ± 0.09 Å, respectively, as superimposed using a full set of the residues described above. The N-terminal myristoyl group, helix A, EF-1, EF-2, EF-3 and EF-4 are shown in purple, pink, green, red, blue and yellow, respectively. Although all the helices and sheets are well defined, the linker between helices H and I, the EF-hand loop regions and the C-terminal 13 residues are less well defined, owing to the lack of a large number of experimental distance and dihedral angle restraints. The r.m.s.d. from the distance restraints is 0.023 ± 0.0004 Å and from dihedral restraints is $0.32 \pm 0.01^\circ$. The average r.m.s.d. values from idealized geometry for bonds, angles and impropers are 0.007 Å, 2.1° and 0.9° , respectively. The total, experimental distance and Lennard-Jones potential energies are $3,983 \pm 15$, 78.3 ± 6.2 and -572 ± 22 kcal mol⁻¹, respectively (calculated with the use of square-well potentials for the experimental distance empirical energy term with a force constant of 50 kcal mol⁻¹ Å⁻²). None of the structures has violations greater than 0.40 Å (for distance restraints) and 3.0° (for dihedral angle restraints). b, Schematic ribbon drawing of the energy-minimized average structure of Ca^{2+} -free, myristoylated recoverin.

Heavy atoms of the N-terminal myristoyl group are shown as a space-filling model. The helices (A, residues 4–16; B, 25–38; C, 46–56; D, 66–73; E, 83–91; F, 98–109; G, 119–132; H, 135–140; I, 148–159; J, 169–178; K, 180–186) are labelled, and the colour coding is the same as in a. The model was generated using QUANTA (Molecular Simulations Inc., Waltham, Massachusetts).

METHODS. Structure calculations were performed with the YASAP protocol²⁵ within X-PLOR²⁶ as previously described²⁷. Structure calculations used 2,710 inter-proton distance restraints (comprising 798 intra-residue, 636 sequential, 460 short-range, 668 long-range and 148 protein-myristate) obtained from heteronuclear three-dimensional nuclear Overhauser effect (NOE) spectra of the sample whose N-terminal myristoyl group and protein are uniformly labelled with ^{13}C , and the sample whose protein portion is labelled with ^{15}N and myristoyl group with ^{13}C . The NOEs were interpreted by using previously reported backbone assignments¹⁴ and recently obtained side-chain assignments, which included the stereospecific assignments of the Leu and Val methyl groups accomplished by using the samples whose Leu or Val methyl groups were stereospecifically labelled with ^{13}C . All of the ^1H and ^{13}C signals of the myristoyl group were assigned by a combination of heteronuclear multidimensional NMR experiments on ^{15}N -labelled recoverin with a ^{13}C -labelled myristoyl group, together with careful inspection of the calculated structures with the use of the distance restraints associated with unambiguously assigned myristoyl signals (that is, C_2H , C_3H and $\text{C}_{11}\text{H}-\text{C}_{14}\text{H}$). NOE-derived distance restraints were classified as 1.8–2.7, 1.8–3.3, 1.8–5.0 and 1.8–6.0 Å (1.8–2.9, 1.8–3.5, 1.8–5.0 and 1.8–6.0 Å for NOEs involving the backbone amide protons) on the basis of the relative NOE cross-peak intensities. In addition to the NOE-derived distance restraints, 140 distance restraints for 70 hydrogen bonds and 238 dihedral angle restraints were included in the structure calculation.

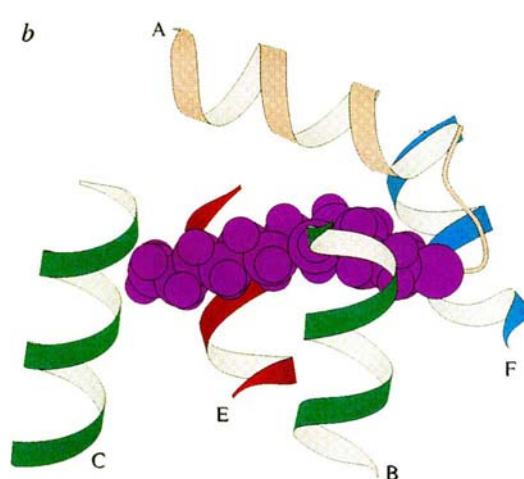
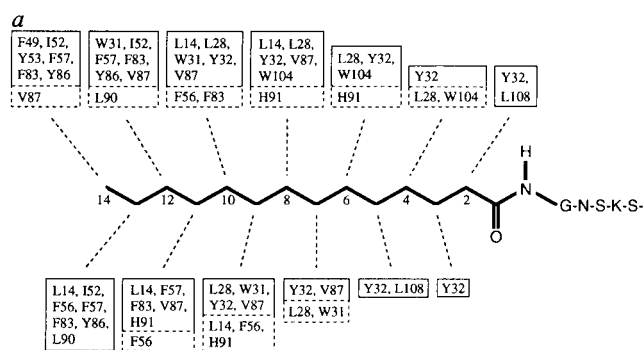


FIG. 2 **a**, Summary of close contacts between the myristoyl group and protein side chains. Only hydrophobic residues within 5 Å of the myristoyl group are shown. The contacts observed by NOEs are shown in solid boxes, otherwise in dotted boxes. Myristoyl pocket residues from five helices: (1) Leu 14 from helix A; (2) Leu 28, Trp 31 and Tyr 32 from helix B; (3) Phe 49, Ile 52, Tyr 53 and Phe 56 from helix C; (4) Phe 83, Tyr 86, Val 87, Leu 90 and His 91 from helix E; (5) Trp 104 and Leu 108 from helix F. **b**, Ribbon diagram showing five helices surrounding a

space-filling model of the myristoyl group. The helices are labelled, and the colour coding is the same as in Fig. 1 except that residues 2 and 3 are in pink. The figure was generated using MOLSCRIPT²⁸.



FIG. 3 Sequence alignment of the recoverin subfamily of EF-hand calcium-binding proteins. The amino acid sequences of bovine recoverin (REC)³, frog S-modulin (SMO)⁵, chicken visinin (VIS)¹⁵, rat hippocalcin (HIP)²⁹, bovine neurocalcin (NEU)³⁰, rat neural visinin-like Ca²⁺-binding protein (NVP)¹⁸ and *Drosophila* frequenin (FRE)³¹ are compared. The N-terminal myristoyl group is indicated as m, at which the residue numbering starts. Secondary structural elements of Ca²⁺-free, myristoylated

recoverin and their notation as well as EF-hand motifs are shown above the sequences: α -helices and β -strands are indicated as double-headed arrows and open arrows, respectively. Asterisks denote the conserved hydrophobic amino acid residues. The hydrophobic residues that are found within 5 Å of the myristoyl moiety are shown in bold. Positively charged residues near the N and C terminus are underlined.

EF-2 is rotated approximately 45° with respect to EF-3 (Fig. 4). A backbone hinge rotation at the domain linker allows a rearrangement of the domain interface such that EF-3 makes contact with either helix E (Ca²⁺-free) or helix D (single-Ca²⁺-bound). The differences between the solution and crystal structures provide possible clues about the molecular mechanism of the Ca²⁺-myristoyl switch.

The myristoyl-binding pocket described here is different from the hydrophobic crevice found in the unmyristoylated protein, which was thought to be the myristoyl or target protein binding site²⁰. Several aromatic and other nonpolar residues lining this crevice are indeed in contact with the myristoyl group in our solution structure. However, the participation of five helices in forming the myristoyl-binding pocket (Fig. 2b), the almost complete burial of the myristoyl group, and the proximity of the carbonyl end of the myristoyl group to EF-3 (Fig. 1b) were not previously envisaged.

Myristoyl switches are used in transducing diverse signals. For example, ADP-ribosylation factors (ARFs) are myristoylated

small GTP-binding proteins that play key roles in vesicular transport. ARFs exhibit GTP-dependent partitioning into membranes, suggesting that ARFs contain a GTP-myristoyl switch²¹. Myristoylated alanine-rich C kinase substrate (MARCKS), nitric oxide synthase and members of the Src tyrosine kinase family bind to membranes, whereas phosphorylated forms of these proteins do not²²⁻²⁴. A phosphoryl-myristoyl switch may couple phosphorylation to a change in accessibility of the myristoyl group. The structure reported here is the first atomic resolution view of a myristoylated protein possessing a signalling switch. The sequestration of the myristoyl group seen in recoverin is likely to be a key element of one of the states of myristoyl switches generally.

Received 10 April; accepted 25 May 1995.

1. Stryer, L. *J. biol. Chem.* **266**, 10711-10714 (1991).
2. Koch, K.-W. *Rev. Physiol. Biochem. Pharmac.* **125**, 149-192 (1994).
3. Dzhoghlev, A. M. et al. *Science* **251**, 915-918 (1991).
4. Gray-Keller, M. P., Polans, A. S., Palczewski, K. & Detwiler, P. B. *Neuron* **10**, 523-531 (1993).

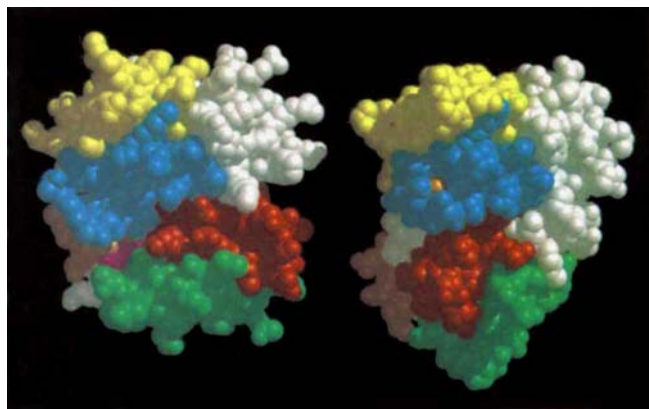


FIG. 4 Comparison of the structures of Ca^{2+} -free, myristoylated recoverin (left, derived from NMR) with that of single- Ca^{2+} -bound unmyristoylated recoverin (right, derived from X-ray²⁰). Both structures are shown as space-filling models. The colour coding is the same as in Fig. 1 except that residues 2 and 3 are in pink and the Ca^{2+} ion is in orange. The C-terminal domains of the two forms were aligned to show the 45° rotation of the N-terminal domains. In the Ca^{2+} -free structure, the hydrophobic residues involved in the domain interface are: Leu 9 and Ile 13 (helix A), Tyr 32 (helix B), His 68, Val 69 and Phe 73 (helix D), Val 87, Ile 88, Ala 89 and His 91 (helix E), Met 92 (linker between helices E and F), Trp 104, Ala 105, Leu 108 and Tyr 109 (helix F), Ile 125, Ala 128, Ile 129 and Met 132 (helix G), and Ile 186 (helix K). In the single- Ca^{2+} -bound structure, the domain interface consists of Phe 57 and Ala 60 (linker between helices C and D), Ala 64, Tyr 65, His 68 and Val 69 (helix D), Phe 73 (between helices D and E), Met 92 (helix E), Trp 104 and Leu 108 (helix F), Val 111 (between helices F and G), Ile 125, Ala 128 and Ile 129 (helix G), Met 132 (linker between helices G and H), and Pro 190 (C terminus). The helices of the crystal structure are named starting from the N terminus. The figure was generated using MIDAS³².

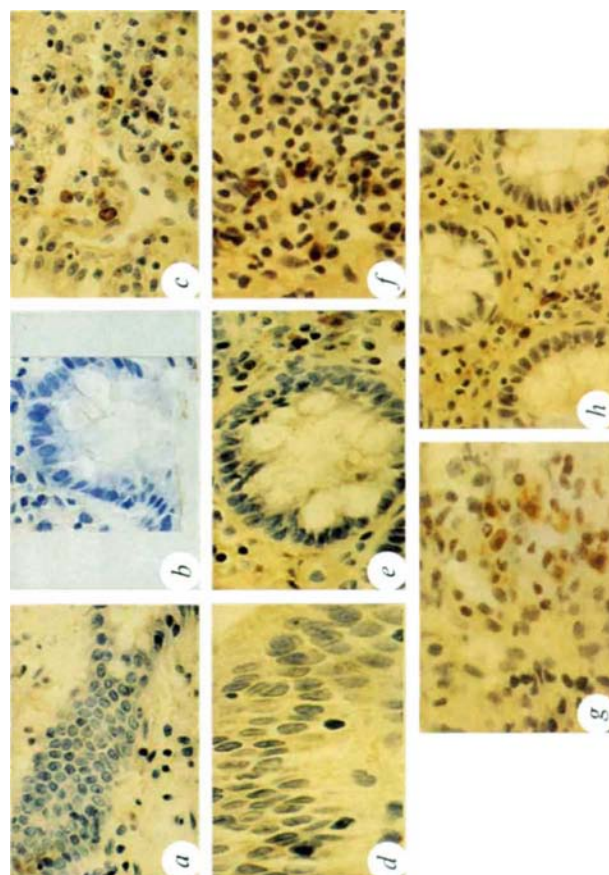
CORRECTION

Tumorigenesis and metastasis of neoplastic Kaposi's sarcoma cell line in immunodeficient mice blocked by a human pregnancy hormone

Yanto Lunardi-Iskandar, Joseph L. Bryant, Robert A. Zeman, Victor H. Lam, Felipe Samaniego, Jacques M. Besnier, Philippe Hermans, Alain R. Thierry, Parkash Gill & Robert C. Gallo

Nature 375, 64–68 (1995)

DUE to a computer mounting error the same four figure parts in Fig. 2C (c, f, g, and h) were inadvertently submitted. The correct Fig. 2C is shown. In the legend to Fig. 2C where it reads '... polyclonal antibody to hCG...' it should read '... polyclonal antibody to purified hCG receptors ...'.



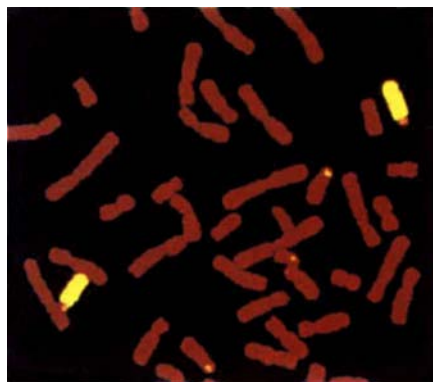
- Kawamura, S., Hisatomi, O., Kayada, S., Tokunaga, F. & Kuo, C.-H. *J. biol. Chem.* **268**, 14579–14582 (1993).
- Dizhoor, A. M. et al. *J. biol. Chem.* **267**, 16033–16036 (1992).
- Ames, J. B., Porumb, T., Tanaka, T., Ikura, M. & Stryer, L. *J. biol. Chem.* **270**, 4526–4533 (1995).
- Zozulya, S. & Stryer, L. *Proc. natn. Acad. Sci. U.S.A.* **89**, 11569–11573 (1992).
- Dizhoor, A. M. et al. *Science* **259**, 829–832 (1993).
- Babu, Y. S., Bugg, C. E. & Cook, W. J. *J. molec. Biol.* **24**, 191–204 (1988).
- Herzberg, O. & James, M. N. G. *J. molec. Biol.* **203**, 761–779 (1988).
- Zheng, J. et al. *Protein Sci.* **2**, 1559–1573 (1993).
- Chow, M. et al. *Nature* **327**, 482–486 (1987).
- Ames, J. B., Tanaka, T., Stryer, L. & Ikura, M. *Biochemistry* **33**, 10743–10753 (1994).
- Yamagata, K., Goto, K., Kuo, C.-H., Kondo, H. & Miki, N. *Neuron* **2**, 469–476 (1990).
- Kobayashi, M., Takamatsu, K., Saitoh, S. & Noguchi, T. *J. biol. Chem.* **268**, 18898–18904 (1993).
- Ladant, D. *J. biol. Chem.* **270**, 3179–3185 (1995).
- Kuno, T. et al. *Biochem. biophys. Res. Commun.* **184**, 1219–1225 (1992).
- Resh, M. D. *Cell* **76**, 411–413 (1994).
- Flaherty, K. M., Zozulya, S., Stryer, L. & McKay, D. B. *Cell* **75**, 709–716 (1993).
- Helms, J. B., Palmer, D. J. & Rothman, J. E. *J. Cell Biol.* **121**, 751–760 (1993).
- Kim, J., Shishido, T., Jiang, X., Aderem, A. & McLaughlin, S. *J. biol. Chem.* **269**, 28214–28219 (1994).
- Michel, T., Li, G. K. & Busconi, L. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6252–6256 (1993).
- Walker, F., deBlaquiere, J. & Burgess, A. W. *J. biol. Chem.* **268**, 19552–19558 (1993).
- Nilges, M., Gronenborn, A. M., Brünger, A. T. & Clore, G. M. *Protein Engng* **2**, 27–38 (1988).
- Brünger, A. T. *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1993).
- Bagby, S., Harvey, T. S., Eagle, S. G., Inouye, S. & Ikura, M. *Structure* **2**, 107–122 (1994).
- Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).
- Kobayashi, M., Takamatsu, K., Saitoh, S., Miura, M. & Noguchi, T. *Biochem. biophys. Res. Commun.* **189**, 511–517 (1992).
- Okazaki, K. et al. *Biochem. biophys. Res. Commun.* **185**, 147–153 (1992).
- Pongs, O. et al. *Neuron* **11**, 15–28 (1993).
- Ferrin, T. E., Huang, C. C., Jarvis, L. E. & Langridge, R. J. *molec. Graph.* **6**, 13–27 (1988).

ACKNOWLEDGEMENTS. We thank M. Kainosho for providing stereospecifically labelled leucine and valine, L. Kay for his expert help in NMR experiments, T. Neubert for mass spectrometry analysis on the myristoylated recoverin samples, F. Delaglio and D. Garrett for providing NMR data processing software, and N. Tjandra and A. Bax for providing a program for inter-helix angle calculation. This work was supported by grants to M.I. from the MRC of Canada and HFSP, and grants to L.S. from NIH. T.T., J.B.A. and T.S.H. are supported by OCI/Amgen, NIH and NATO postdoctoral fellowships, respectively.

What's in a label?

For labelling technology, here is the 'short list' of new products — a radiolabelled DNA purification kit, reagents for analysis of cytokine reagents, a DIG detection ELISA kit and a line of streptavidin-coated microplates.

Cambio has added **StarFISH mouse chromosome paints** to its range of products (*Reader Service No. 100*). The paints are designed to offer easy detection of translocation and trisomy in mouse chromosomes. Labeled with biotin, they are available for all individual chromosomes, as well as certain combinations. All paints are highly chromosome-specific and are available for



StarFISH chromosome painting system.

each and every individual human chromosome. A variety of different labels are available for human paints, namely FITC, biotin and cyanine 3, a label that is inherently bright and requires no further amplification. The system also allows the detection of translocations and insertions on metaphase chromosomes. Marker chromosomes can also be identified and complex karyotypes analysed. The company has also developed a number of visualization kits including dual-colour paint kits, which enable several different chromosomes to be visualized simultaneously.

Oncor has developed the **ApopTag Plus kit** for *in situ* staining and detection of apoptotic cells in the context of surrounding tissue (*Reader Service No. 101*). This product is available in peroxidase and fluorescein versions and includes two control slides, extra



Oncor ApopTag Plus detection kit.

reaction buffer, DAB substrate and easy access vials. The kit is designed to detect single apoptotic cells and is said to be an improvement over *in situ* polymerase translation in distinguishing apoptosis from necrosis.

For **purification of radiolabelled DNA** there is the new, large GeneClean kit now available from Stratech Scientific (*Reader Service No. 102*). GeneClean III can perform up to 300 preps and can be used with TAE- or TBE-buffered agarose gels. It comes complete with LabelBlock solution, which enables purification of radiolabelled DNA. This allows isolation from fossils, museum specimens and bones. Finally, the new spin module can be used with all the silica-based extraction kits for cleaner DNA/RNA with no matrix carry over and consistent yield. Spin module evaluation samples and protocols are available on request.

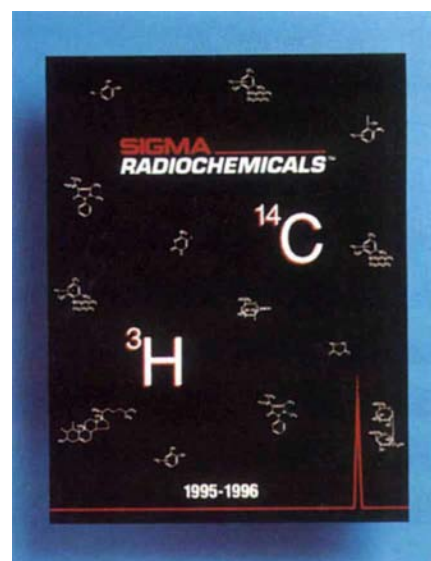
Radiolabelling

Wallac offers a **confidential custom labelling service** to researchers in academia and industry (*Reader Service No. 103*). The service includes the labelling of a wide range of biological compounds, from large proteins to small molecular weight molecules, such as oligonucleotides and haptens. Additionally, the company can custom-coat microtitre plates with antibodies, antigens and other proteins. Also offered with this service is optimization of assays for the company's time-resolved fluorometry system, Delfia. This system is designed to be a sensitive, fast and non-isotopic alternative to radioisotopic ELISA. The non-radioactive technology also allows the measurement of more than one analyte in a single well. Because of the proprietary nature of certain compounds that may be requested, the company offers its labelling service to customers under mutual non-disclosure agreements. Labelled compounds and coated plates can be provided in small quantities as well as in bulk.

Based on the existing **tritium labelling** service, LabellingPlus from DuPont NEN is designed to offer researchers a fast, economical alternative to custom synthesis with a finished product in approximately four weeks and a fixed price, for easier budgeting (*Reader Service No. 104*). However, the product also includes specific and total activity determination, stated to be greater than 90 per cent, provision of a technical data sheet and comparison of the product against a cold standard. Although the service is likely to be used primarily for the labelling

of ligands and non-peptide drugs, it can also be used to label peptides, steroids and lipids.

Sigma also offers the **1995-96 radiochemicals catalogue**, which contains a comprehensive catalogue list of over four hundred ^{14}C and ^3H compounds ready for immediate shipment (*Reader Service No. 105*). New products have been added to complement the



Sigma's buyer's guide to Radiochemicals.

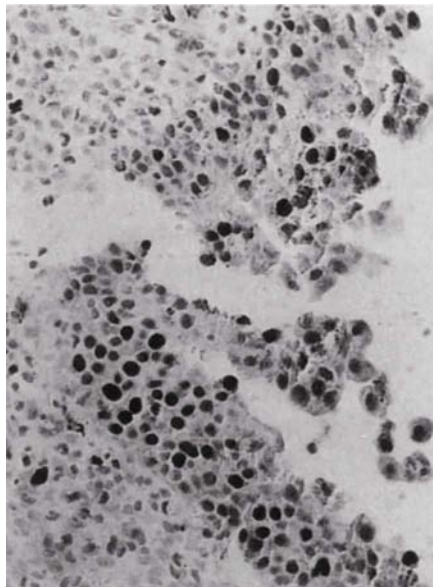
company's existing lines of labelled amino acids, carbohydrates, enzymes, molecular weight markers, nucleosides, nucleotides, organic chemicals and proteins. New indices for tritiated products, and for all products arranged by chemical group, have been added. Also included are liquid scintillation cocktails and an extensive radiochemical safety line.

Antibodies

R&D Systems' Fluorokine kits contain biotinylated cytokines and avidin-FITC secondary reagents for the **analysis of cytokine receptors by flow cytometry** (*Reader Service No. 106*). The biotinylated cytokine kits have recently been upgraded to give a brighter signal. Kits currently available include human IL-1 α , IL-2, IL-6, IL-10, LIF, MIP-1 α and SCF. Included in the kits are a protein negative control and an antibody specificity control. The kits are designed to be easy to use — cells of interest are incubated with biotinylated cytokine, secondary reagent is added, the cells are washed and are then ready for flow analysis. Biotinylated cytokines may offer a more complete picture than antibodies for studying cytokine recep-

tors as cytokines can bind to several receptor components whereas an antibody may only recognize one receptor component.

Oncogene Science has released two new **antibodies against the Rb tumour suppressor gene protein (p105)** (*Reader Service No. 107*). Rb (Ab-5) may be used for



Rb (Ab-5) staining of formalin-fixed, paraffin-embedded normal human bladder after thermal pretreatment.

paraffin section staining and immunoblotting. Rb (Ab-6) will react with the Rb protein from a wide range of species, including human and *Xenopus* by immunoblotting.

ELISAs

Antigenix has announced the availability of an **ELISA kit for the measurement of human monocyte chemotactic and activating factor (MCAF)**, a recently identified chemokine (*Reader Service No. 108*). Also known as MCP-1, this chemokine is associated with the pathogenesis of pulmonary fibrosis, rheumatoid arthritis, atherosclerosis and allergy. The research kit is based on a pre-coated



Antigenex ELISA kit for human MCAF.

microplate and contains dual sets of standards for multiple assays of human serum, plasma, or cell culture supernatant. The sensitivity is stated to be 2.0 pg ml^{-1} . A similar ELISA kit for human IL-8 is also available.

Completing a range of **kits to detect digoxigenin (DIG) and biotin labelled biomolecules** is the DIG detection ELISA (fluorescence) kit available from Boehringer Mannheim (*Reader Service No. 109*). The kit has been designed for enhanced sensitivity using the new AttoPhos substrate system which has a large Stoke's shift (140 nm), low background fluorescence and a large dynamic measurement range. The manufacturer states that this increases the sensitivity by a factor of 100 compared to the more commonly used 4-methyl-umbelliferyl-phosphate fluorescent substrate. Streptavidin-coated microplates are said to give high binding capacity, high coating homogeneity and high intra- and inter-assay reproducibility with no leaching. The kit is supplied with a standardized control peptide containing one molecule of digoxigenin and one molecule of biotin. This allows the signal to be correlated to molar values and may be used to ensure reproducibility of the results.

Odds and ends

To complete a range of high-quality microplates and strips, Labsystems has developed **streptavidin-coated microplates** (*Reader Service No. 110*). The plates offer a universal solid-phase for immunoassays, DNA hybridization assays, and receptor and cell studies. The covalent bonding of streptavidin to the plate surface is designed to ensure stability, even under demanding assay conditions. Streptavidin solid phase is extremely useful for immobilizing small or hydrophilic molecules as adsorption properties of these molecules are usually poor. Compounds such as DNA, peptides, polysaccharides and haptens can be effectively immobilized on a streptavidin-coated solid phase as their biotin derivatives. DNA hybridization assays using coated microplates offer several advantages over traditional gel methods, including improved sensitivity and specificity, and most importantly a shorter, simpler assay that is easy to automate.

Sigma Biosciences has released Sigma ImmunoNotes number 13, which discusses **fluorescent *in situ* hybridization using Cy3** (*Reader Service No. 111*). The article explains the procedure and the advantages of using the company's streptavidin-labelled Cy3 reactive dye. New products listed include anti-digoxin, ExtrAvidin, anti-biotin reagents, anti-glutathione-S-transferase (GST) and a monoclonal anti-VSV glycoprotein.

SuperMount is an **aqueous-based permanent mounting medium** developed by BioGenex for the slide-mounting of all biological specimens (*Reader Service No. 112*). This product is suitable for use with most chromogens commonly used in immunohistochemistry and *in situ* hybridization. Specifically designed to meet the mounting needs of the histological labora-

tory, SuperMount can be used for the permanent mounting of all biological specimens for microscopic evaluation and indefinite storage, including stained tissue sections, cytospin preparations and blood smears. When applied in a thin coat and allowed to harden, it forms a highly transparent coating that permanently preserves stained tissue sections, making it a suitable choice for mounting sections to be photographed. This product is compatible with most aqueous and organic soluble dyes, including AEC, DAB, Fast Red, new Fuchsin and BCIP/NBT. It is also compatible with specimens stained with fluorescent dyes, including rhodamine, fluorescein, Texas Red and phycoerythrin and its related derivatives, such as phycocyanin and allophycocyanin.

Celsis has announced the introduction of the new Optocomp series of **luminometers** (*Reader Service No. 113*). Available in two models that can be programmed to deliver



Celsis' Optocomp II luminometer.

up to two reagents with a coefficient variance of less than one per cent using two, one or no injectors with variable delay and counting times. These products can provide quality control and quality assurance personnel with a new tool for rapid microbial detection. □

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENT

OLIGONUCLEOTIDES

Synthesised, purified and shipped in 3 days

£2.50 per base

(All inclusive)

0.2 µM scale (approx. 500 µg)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit
 King's College School of Medicine & Dentistry
 Tel: 071 346 3126 Fax: 071-733 3877

Reader Service No.10